NEUROIMAGING INVESTIGATIONS OF LANGUAGE TO AID PAEDIATRIC NEUROSURGICAL DECISION MAKING

Louise Jane Croft
UCL

Thesis submitted for the degree of

DOCTOR OF PHILOSOPHY

I, Louise Jane Croft, confirm the work presented in this thesis is my own. Where information has been derived from other sources, I confirm this has been indicated in the thesis.

Signed: ……..

Date: ……..

“*Just keep swimming”*

− Dory −

ABSTRACT

Childhood onset epilepsy can have a profound effect on cognitive, emotional and behavioural development. As such, early intervention is crucial. Approximately 25-50% of children with epilepsy show resistance to medication however. For these children, neurosurgical intervention may be considered. The decision for surgery is a multi-disciplinary process, including functional Magnetic Resonance Imaging (fMRI) to assess the risk posed by surgery to language.

Based on feasibility and behavioural pilot studies, I developed an fMRI task panel optimised for pre-surgical investigations of language in children. This task panel maps different language systems (word retrieval, sentence generation, auditory comprehension and reading comprehension) and localises critical language functions (semantic and syntactic processing). I validated this task panel in healthy children (N=43, 5-16 years). This included assessments of scan quality, comparison of methods for artefact repair, and definition of typical activation patterns. I also piloted the new task panel in a representative sample of children with epilepsy, who were being considered for surgery (N=13, aged 5-16 years). Patient case studies are reported to highlight methodological challenges associated with localisation of critical language regions on an individual basis.

Finally, I present experimental analyses which highlight the importance of the ventral system to semantic processing. Activation in this network was reduced in children with epilepsy and predicted language outcome. Further investigation showed prolonged development of specific nodes within this system, supporting multimodal semantic processing (independent of effort and performance accuracy). These regions included ventral occipito-temporal cortex, whose role in semantic processing has so far been underappreciated in the developmental literature. These analyses provide evidence for a core language system, which may be crucial for post-surgical language outcome.

The findings from this thesis contribute towards extending and improving the role of fMRI in the surgical decision-making process, with the potential for improving long term outcome. They also contribute to models of typical and atypical language development.

ACKNOWLEDGEMENTS

This thesis was made possible by a great many people, whom I would like to thank. Firstly I would like to thank Torsten Baldeweg, for teaching me how to write, for always challenging me, and for all the long hours he sat with me, helping me understand things better. Secondly, Peter Rankin, for setting me on this path and for continued inspiration. I would also like to send a heartfelt thank you across the Atlantic, to Dr Gaillard and Madison Berl - my adopted supervisors - for taking me under their wing, and for investing so much in my professional development, when they didn't have to. In addition to these people, I would like to thank everyone in the Developmental Cognitive Neurosciences Department for providing a constant source of support and encouragement throughout the years, and especially the difficult last 6 months.

I also have to thank all of my collaborators, in particular: Prof Cathy Price, Dr Tom Hope, Dr ʻŌiwi Parker Jones, and Suz Prejawa, for sharing their stimuli and their time, and in particular Tom and ʻŌiwi, for their technical support; all those working on the Epilepsy Surgery Programme at GOSH, particularly Prof Helen Cross, Dr Kelly St Pier and Dr Maria Centeno; Prof Stuart Rosen for helping me make the aliens; Dr David Carmichael and Tim Tierney for helping me understand artefacts; Dr Marko Wilke, for providing technical help; Lauren Zimmaro, Leigh Sepeta and Ben Yerys, for helping me during my time in Washington. I am also grateful to the UCL Charlotte and Yule Bogue Research Fellowship in honour of Sir Charles Lovatt Evans and A.J. Clark, without which Chapter 7 would not have been possible.

I would also like to extend my thanks to all the families who took part in this research, who gave up so much time. Particularly I would like to thank the children from Great Ormond Street Hospital, who chose to take part in this research at a challenging time in their lives.

Finally, I would like to thank my wonderful family and friends, who have supported me over the last few years, especially my parents; for always encouraging me to follow what I enjoy most, and for guiding me. Finally, Dom, I would like to thank you. This thesis is as much a product of your patience, support and optimism, as it is my own sweat and tears. You got me through this and I owe it to you.

PUBLICATIONS RESULTING FROM THIS THESIS

Chapter 1 is based on published content (Croft et al., 2013). Where appropriate, discussion and analyses have been amended for this thesis. All writing, analyses and figures are my own. Data were collected as part of a previous study, under the supervision of Dr Torsten Baldeweg and Dr Peter Rankin.

Chapter 7 is based on content accepted for publication (Croft et al., in press). Where appropriate, discussion and analyses have been amended for this thesis. All writing, analyses and figures are my own. Data were collected as part of a previous study performed at the Children's National Medical Centre, Washington DC., under the supervision of Dr Madison Berl and Dr William Gaillard, made possible by the UCL Charlotte and Yule Bogue Research Fellowship in honour of Sir Charles Lovatt Evans and A.J. Clark.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

How can neuroimaging investigations of language inform the neurosurgical decision making process for children with epilepsy?

In this introductory chapter, I first define drug-resistant epilepsy and describe its effects on cognitive development. I introduce surgery as an alternative treatment option for children with drug resistant epilepsy, and discuss the impact of surgery on long term seizure and cognitive outcome. I describe the multi-disciplinary decision-making process aimed at assessing the risk posed by surgery to cognition, and highlight three stages of this process which can be informed by functional magnetic resonance imaging (fMRI) investigations of language. By reviewing the current literature, I highlight areas where the contribution of language fMRI to the neurosurgical decision-making process can be improved. These become the focus of this thesis.

1. NEUROSURGICAL INTERVENTION FOR DRUG RESISTANT EPILEPSY

Epilepsy is a neurological condition associated with abnormal communication between neurons, which causes seizures. Approximately 50 million people in the world have epilepsy (WHO, 2001), including ~1 in 220 children in the UK (JEC, 2010). Epilepsy is not a single condition; there are at least 28 possible diagnoses for childhood onset epilepsy, at least 14 types of seizure according to recent classification criteria, and the cause may be genetic, structural-metabolic or unknown (Berg et al., 2010).

Drug resistant epilepsy is defined by the failure of two appropriately chosen and adequately trialled and tolerated antiepileptic drugs to achieve sustained seizure freedom (Kwan et al., 2010). Approximately 25-30% of people with epilepsy show resistance to medication (Dichter & Brodie, 1996; Picot, Baldy-Moulinier, Daures, Dujols, & Crespel, 2008). Children with drug resistant epilepsy are vulnerable to cognitive impairment, behavioural problems and poor academic attainment (Berg, 2011; Cormack et al., 2007; Parkinson, 2002). These problems accumulate over time, and can be exacerbated by on-going seizure activity and antiepileptic medication (Bjornaes, Stabell, Henriksen, & Loyning, 2001; Caplan et al., 2009). By adulthood, psychiatric comorbidities and intellectual disability are more frequent in drug resistant patients,

compared to those with controlled epilepsy (Picot et al., 2008). Early intervention is therefore crucial, and neurosurgery may be considered as an alternative when medication fails.

Surgery aims to reduce the impact of seizures on a child's quality of life by removing the epileptogenic area of the brain. Seizure freedom can be achieved post-operatively in 48-86% of cases, depending on the surgical procedure (Lettori et al., 2008; Pulsifer et al., 2004; Skirrow et al., 2011; Tellez-Zenteno, Hernandez Ronquillo, Moien-Afshari, & Wiebe, 2010; Terra-Bustamante et al., 2005; Wyllie et al., 1998). By reducing seizure frequency and the need for anti-epileptic medication, surgical intervention also holds the capacity to enhance the patient's developmental potential (Skirrow et al., 2011), especially when intervention occurs early (Loddenkemper et al., 2007; Spencer & Huh, 2008).

To establish whether a patient is suitable for neurosurgical intervention, multi-disciplinary assessments are performed; 1) to determine whether the epileptogenic area can be localised (thus providing a surgical target), and 2) to evaluate the risk of cognitive deficits post-operatively (Hirsch & Arzimanoglou, 2004). For the latter, neuroimaging investigations of language are crucial.

Almost 50% of neurosurgical procedures in children with drug resistant epilepsy are performed within perisylvian regions (Harvey, Cross, Shinnar, Mathern, & Taskforce, 2008). Language is therefore particularly vulnerable to surgical resection, and up to 34% of patients show post-operative declines in verbal IQ, naming and language comprehension (Sherman et al., 2011). Functional magnetic resonance imaging (fMRI) and intracranial mapping (ICM) aim to assess and minimize the risk posed by neurosurgery to motor and language functions (primarily their proximity to the surgical target). Consequently, for a large proportion of epilepsy patients neuroimaging is extremely important, to guide treatment and improve quality of life. This introductory chapter will focus on the application of fMRI for assessing and minimising the risk posed by surgery to language for children with drug resistant epilepsy.

2. THE ROLE OF LANGUAGE FMRI IN THE PRE-SURGICAL DECISION-MAKING PROCESS

Language fMRI can inform several stages of the neurosurgical decision-making process, outlined below. In the following section I go on to review how well fMRI currently informs each of these stages, to highlight areas for improvement which form the basis of the aims for this thesis.

2.1 Providing a model of typical development

In the healthy population, fMRI informs models of typical language development; a necessary baseline to which patients can be compared, and from which predictions can be made. This is especially important for paediatric patients, when predictions are made in the transitory context of development, and are further complicated by brain plasticity and the potential for functional reorganisation.

2.2 Assessing language lateralisation

The capacity for functional re-organisation elevates the prevalence of atypical (right lateralised or bilateral) language in epilepsy patients, which is harder to predict based on clinical factors (Gaillard et al., 2007). Language fMRI can assess the extent of functional re-organisation, and is therefore crucial to determine whether a patient has left lateralised, right lateralised or bilateral language representation. Measures of language laterality play an early role in the decision making process, helping to establish whether the patient is suitable to be considered as a surgical candidate. If language lateralises to the unaffected hemisphere it is unlikely to present an obstacle to the decision for surgery. If the planned resection involves the language dominant hemisphere however (especially perisylvian regions) intrahemispheric language mapping is required, to establish the risk posed by surgery to language.

2.3 Localising critical language sites

When surgery involves the language dominant hemisphere, intrahemispheric language mapping is used to localise eloquent cortex which must be preserved during surgery. This information can influence outcome, by guiding the location and extent of resection (Korkman et al., 2005; Schwartz, 2009; Schwartz, Devinsky, Doyle, & Perrine, 1998). For children with drug resistant epilepsy, the surgical targets

(structural lesions) most often involve left perisylvian language regions (Figure 1), and the majority of focal surgical procedures (~23%) involve the temporal lobes (Harvey et al., 2008). Anterior temporal resections are performed in ~80% of children with temporal surgical targets (Clusmann et al., 2004). This procedure removes anterior and ventral temporal areas and is associated with post-operative naming decline in adults (Davies et al., 1998; Davies, Risse, & Gates, 2005). Mapping language in anterior and ventral temporal regions is therefore crucial for a large majority of surgical candidates.

Localisation of critical language sites in proximity to the surgical target is performed using intracranial mapping (ICM); an invasive technique which provides precise mapping of critical language loci when the margin for error is low (around the borders of a surgical target for example). However, this procedure is unpleasant for the child, time consuming and costly. It is also restricted by the size and placement of the subdural electrode grids. Thus, the development of non-invasive fMRI tasks which can localise language, and are validated with ICM, would be beneficial.

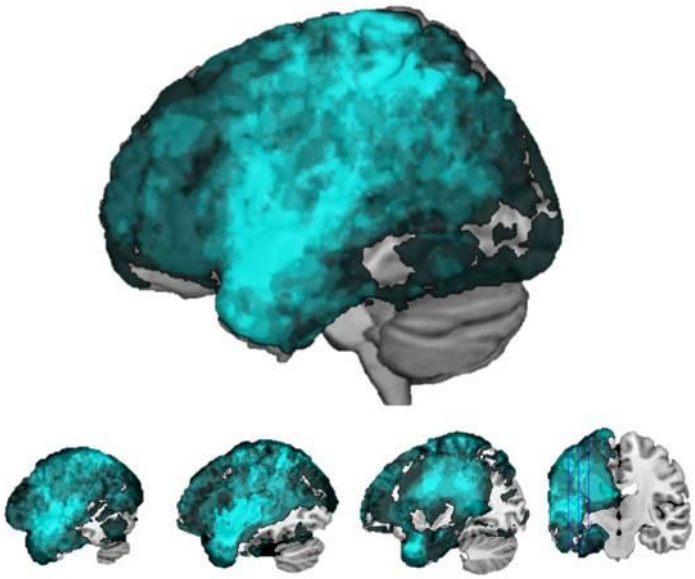


Figure 1: Target regions for neurosurgical resection in children and adolescents with drug resistant epilepsy. Structural lesions are often the target for epilepsy surgery. This lesion overlay map shows the degree of overlap for structural abnormalities in the left hemisphere for a representative sample of surgical candidates. Lighter colours indicate greater overlap between subjects. Lesion data are taken from children and adolescents with drug-resistant epilepsy referred for language fMRI as part of the pre-surgical decision-making process at Great Ormond Street Hospital between 2009 and 2011 (N=52, 25 females, mean age = 12.5 +/-2.7 years, mean age of onset of epilepsy = 5.5 +/-3.8). This figure was created using the npm extension for MRIcron (http://www.mccauslandcenter.sc.edu/mricro/mricro/index.html).

3. TO WHAT EXTENT DOES FMRI CURRENTLY FULFIL THESE ROLES?

I have highlighted three primary roles for language fMRI in the pre-surgical decision making process, including: 1) providing an understanding of typical language development, for use as a baseline in the pre-surgical setting, 2) assessing language lateralisation, and 3) localising language cortex. Here I review the extent to which fMRI currently fulfils each of these roles, with the aim of identifying areas of strength and – crucially - areas for improvement (which will be the focus of this thesis).

3.1 Providing a model of typical development

3.1.1 The mature language network

The main findings of reviews summarising this vast literature, and key studies, are summarised below (Binder, Desai, Graves, & Conant, 2009; Bookheimer, 2002; Poldrack et al., 1999; Price, 2010; Price, 2012; Vigneau et al., 2006). In contrast to the modular Broca-Wernicke model of expressive and receptive language from the 19th century (Broca, 1865; Wernicke, 1874), language is now conceived as a unified network, consisting of distinguishable - yet integrated - functional and anatomical sub-components. The semantic, syntactic, phonological, orthographic, executive and working memory components of a language task modulate the weighting of regional activation throughout this system (Vigneau et al., 2006). Despite this complexity, the language network can theoretically be divided into two parts; an action component (the dorsal pathway) and a perception component (the ventral pathway), similar to the visual system (Hickok, 2009; Hickok & Poeppel, 2000, 2004, 2007; Poeppel, Emmorey, Hickok, & Pylkkanen, 2012; Poeppel & Hickok, 2004; Saur et al., 2008) (Figure 2).

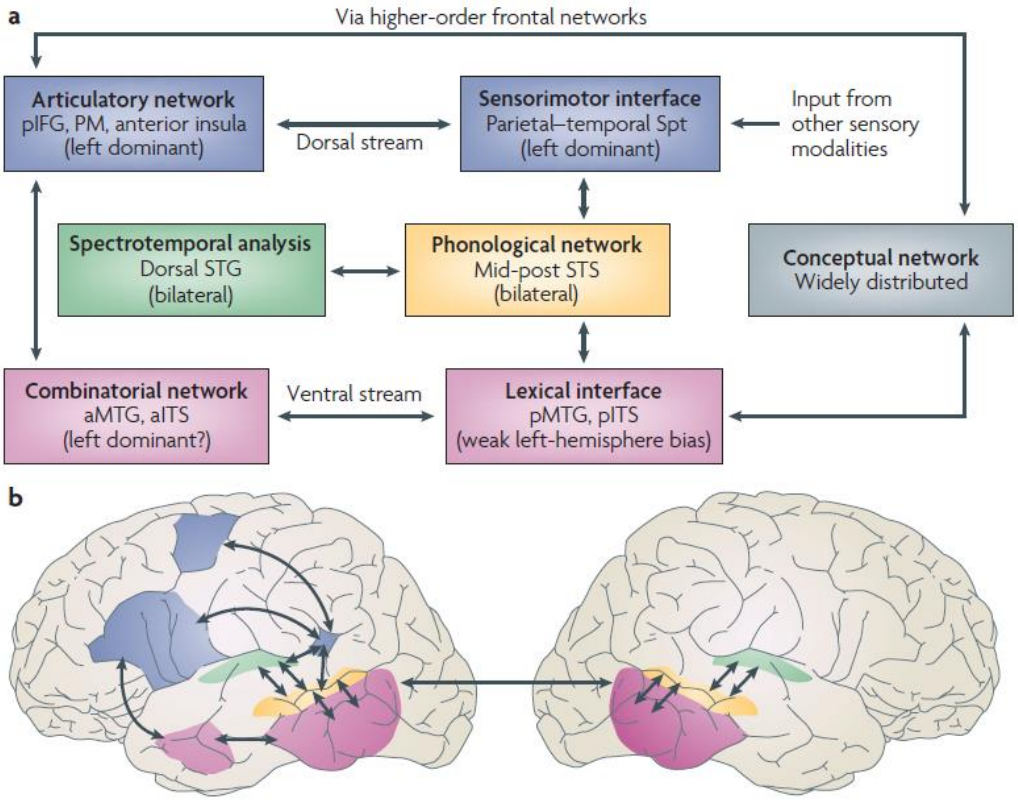


Figure 2: The Dual Stream model of language (*from Hickok and Poeppel, 2007*).

3.1.1.1 The dorsal system

The dorsal pathway engages the primary auditory cortex (Heschl's gyrus; HG), posterior temporal lobe (superior temporal sulcus and planum temporal), inferior parietal lobe (IPL), precentral gyrus, operculum and Pars Opercularis (Brodmann area 44) in the inferior frontal gyrus. These regions are interconnected via the superior longitudinal and arcuate fasciculi (Catani, Jones, & ffytche, 2005; Catani & Mesulam, 2008; Maldonado, Moritz-Gasser, & Duffau, 2011). Within this system sensory inputs (e.g. heard words) are associated with motor signals (e.g. mouth movements) (Blank, Scott, Murphy, Warburton, & Wise, 2002; Price & MacSweeney, 2011; Pulvermüller & Fadiga, 2010). These associations become established during language acquisition via the process of hearing words, reproducing them and receiving sensory feedback. The Pars Opercularis (part of classic Broca's area) is regarded as a crucial component for this process (Price & MacSweeney, 2011).

3.1.1.2 The ventral system

The ventral language pathway establishes a sense of meaning, by integrating external or internal sensori-motor representations of a word with internal information (memories and emotions). This network involves the IPL, temporal lobes (particularly ventro-lateral and anterior regions) and the anterior inferior frontal gyrus (Pars Orbitalis; Brodmann area 47). The anterior temporal lobes (bilaterally) are considered critical to this network, as an amodal semantic hub (Adhami et al., 2006; Binder et al., 2009; Hodges, Patterson, Oxbury, & Funnell, 1992; Holland & Lambon Ralph, 2010; Visser, Jefferies, & Lambon Ralph, 2010), while the Pars Orbitalis exerts constraints during conceptual processing. Several white matter tracts have been implicated in the ventral system, including: the extreme capsule, uncinate fasciculus, middle longitudinal fasciculus and inferior frontal occipital fasciculus (Makris & Pandya, 2009; Makris et al., 2009; Makris et al., 2007; Schmahmann & Pandya, 2007a, 2007b; Schmahmann et al., 2007).

3.1.1.3 An integrated parallel processing loop

Due to the relevance of sensori-motor processing and classic Broca's area to speech production, the dorsal pathway has traditionally been considered the major language pathway, and has specifically been associated with expressive language (Catani et al., 2005). Conversely, the ventral pathway has been associated with receptive language due to its role in semantic processing. However, the internal sensori-motor, memory and affective representations of words become strongly associated over time and due to extensive experience with language. These associations firmly integrate the dorsal and ventral pathways, meaning both are involved in the production and perception of language (Pulvermüller & Fadiga, 2010; Rolheiser, Stamatakis, & Tyler, 2011), forming a loop system as opposed to two separate streams (Weiller, Musso, Rijntjes, & Saur, 2009) (see Figure 3). The middle temporal gyrus (MTG) is a prospective substrate for integration of the two pathways due to its extensive anatomical connectivity, and may be a critical language hub (Turken & Dronkers, 2011).

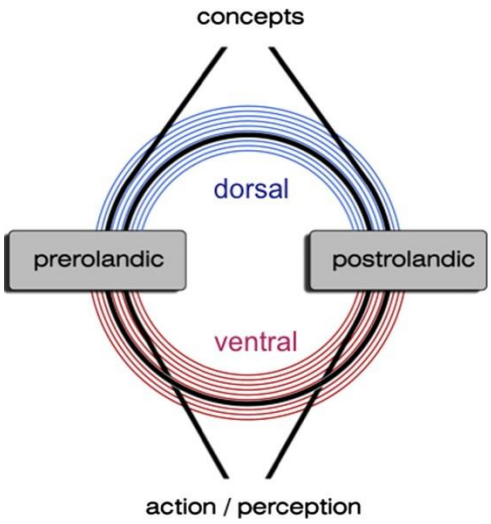


Figure 3: A dual loop system for processing naturalistic language presented by Weiller et al., 2011.

3.1.2 The developing language network

How the language network develops from birth, to create the dual loop system seen in adults is not yet fully understood. However several studies have begun to investigate age-related changes in the language system using fMRI, with the aim of identifying how and when the system becomes 'adult-like'. I review these studies below, and highlight areas where further work is required to build a more comprehensive model of typical language development. I also highlight methodological issues which have so far limited our understanding of language development.

3.1.2.1 Receptive language (comprehension)

Children as young as 4 years of age recruit a left lateralised fronto-temporal network for auditory comprehension, including the left superior temporal sulcus (STS), left MTG, left inferior frontal gyrus (IFG) and right cerebellum (Ahmad, Balsamo, Sachs, Xu, & Gaillard, 2003; Balsamo, Xu, & Gaillard, 2006; Berl et al., 2010; Vannest et al., 2009; Yeatman, Ben-Shachar, Glover, & Feldman, 2010). Activity within frontal regions during auditory comprehension is variable across development, but demonstrates modest increases with age (Berl et al., 2010; Chou et al., 2006). Reading comprehension involves similar regions to auditory comprehension, but generally engages a broader and more bilateral network, including the superior frontal gyrus (SFG), right IFG, lingual gyri and occipital lobes (Balsamo et al., 2006; Gaillard, Balsamo, Ibrahim, Sachs, & Xu, 2003; Gaillard, Pugliese, et al., 2001). These differences are considered to reflect increased task difficulty (reading is a later acquired skill), as well as orthographic processing. The passive comprehension tasks designed for children often fail to identify significant activation in key semantic regions (the anterior and ventral temporal lobes). As such, our understanding of how the ventral pathway develops is currently limited. Higher-level contrasts which localise semantic processing during comprehension may improve mapping of these regions (Binder et al., 2011; Visser et al., 2010).

3.1.2.2 Expressive language (word generation)

Expressive language development has typically been investigated using single word generation tasks, involving single word comprehension, semantic association and word retrieval (Petersen, Fox, Posner, Mintun, & Raichle, 1989). These tasks induce strong and consistently left lateralised activation in children as young as 5 years of age in regions similar to the dorsal pathway described in adults, including: classic Broca's area (IFG) and Wernicke's area (posterior superior temporal gyrus; pSTG), the middle frontal gyrus (MFG), dorsolateral prefrontal cortex (DLPFC), cingulate cortex, supplementary motor area (SMA), inferior temporal gyrus (less consistently), IPL and right cerebellum (Gaillard et al., 2000; Holland et al., 2001; Karunanayaka et al., 2010; Lee et al., 1999; Szaflarski, Holland, Schmithorst, & Byars, 2006; Szaflarski, Schmithorst, et al., 2006; Vannest et al., 2010; Wilke et al., 2005; Wood et al., 2004). With age, these regions become increasingly active (Szaflarski, Holland, et al., 2006; Szaflarski, Schmithorst, et al., 2006; Vannest et al., 2010) and left lateralised (Everts et al., 2009; Holland et al., 2001; Karunanayaka et al., 2010; Szaflarski, Holland, et al., 2006; Szaflarski, Schmithorst, et al., 2006), and facilitate performance (Balsamo et al., 2006; Gaillard et al., 2000; Vannest et al., 2010; Yeatman et al., 2010). However, the effects of age, performance and effort on brain activation during childhood cannot be discerned from these studies due to a lack of in-scanner performance monitoring, and a one-size-fits-all approach to task design. Increasing involvement of left frontal regions in these studies may therefore reflect linguistic or non-linguistic processes (such as executive functioning). Developmentally appropriate language tasks aim to control for age-related differences in effort (Gaillard, Grandin, & Xu, 2001), and suggest core language regions are actually in place early in childhood (Gaillard et al., 2000; Lee et al., 1999; Wood et al., 2004). However few studies have adopted this approach. Many studies have incorporated a button-press response mechanism into language tasks, as a means to monitor and control for task performance. However this introduces a decision element to the task which may modulate activation in frontal regions (this is discussed further in the section below).

Finally, current models of expressive language development are also limited by a focus on single-word generation. In conversation meaning is conveyed through the retrieval and structuring of multiple words in a sentence, according to both semantics and syntax. Single-word generation tasks may therefore underestimate the amount of conceptual processing – and therefore participation of the ventral pathway - in expressive language. Sentence formulation has not been investigated using fMRI in children, despite its ecological validity. This is even more surprising when one considers the protracted development of syntax, which suggests sentence formulation tasks may provide a more sensitive measure for developmental change in the language network.

3.1.2.3 Semantic / syntactic decision making

Decision-based receptive and expressive language tasks have been developed to monitor performance, typically by a forced choice button press response (for example, the *Native American Animal task* reported by Wilke et al., 2006). Compared to passive tasks, these active tasks are associated with increased activity in the left frontal regions (Adank & Devlin, 2010; Cabeza & Nyberg, 2000; Chou et al., 2006; Gaillard et al., 2004; Pulvermüller & Fadiga, 2010; Vannest et al., 2009; Wilke et al., 2005). However, because decision-making increases demands on non-linguistic processes, increased activation in dorsal regions may reflect the development of task-related executive skills instead, such as working memory (Klingberg, Forssberg, & Westerberg, 2002), attention (Adleman et al., 2002) and top-down control (Bitan et al., 2006; Casey, Galvan, & Hare, 2005). To identify changes in the perisylvian network which are independent from top-down control, tasks require a response mechanism which minimises executive load or decision-making, such as overt speech. There are methodological challenges associated with the use of overt speech during fMRI scanning (Barch et al., 1999). However, several methods have been developed to repair movement-related imaging artefacts, which may enable investigations of expressive language development using overt speech (Diedrichsen & Shadmehr, 2005; Friston, Williams, Howard, Frackowiak, &

Turner, 1996; Lemieux, Salek-Haddadi, Lund, Laufs, & Carmichael, 2007; Wilke, 2012).

3.1.3 Summary: An accurate model of typical language development?

Overall the literature suggests children engage similar regions to adults during performance of receptive and expressive language tasks, although the contribution of the ventral system to expressive language may currently be underestimated by these studies, due to a focus on single-word generation. How these systems change with age is also not clear, as the majority of findings are confounded by age-related differences in task performance accuracy and effort. Further, many studies used decision-based language tasks or relatively low-level baseline conditions, which fail to localise age-related changes specific to the language system. In contrast, fMRI studies in adults have shown the value of using high-level contrasts, which have enabled differentiation of sub-networks within the language system (dorsal and ventral) and specific structure-function mappings (Price, 2012). The extent to which models of typical language development can inform more complex surgical decisions in children will remain limited until a more comprehensive understanding of how specific structure-function relationships emerge and change across the course of development has been achieved.

3.2 Assessing language lateralisation

The first assessments of language dominance were performed in the 1940's, using local anaesthesia (Gardner, 1941), and later using the intracarotid amobarbital procedure (IAP) or Wada test (Wada, 1949). During the Wada procedure, sodium amobarbital is injected into one of the internal carotid arteries to shut down function in one cerebral hemisphere. Language functions are then tested, while the patient is awake, and while only one hemisphere is functional. This is then repeated, to assess language functions when the other hemisphere is functioning. This invasive procedure is extremely stressful for children, and is difficult to perform in patients with cognitive impairment and behavioural co-morbidities; although it can be done with varying success (Jansen et al., 2002).

Functional MRI provides a reliable non-invasive alternative to the Wada test. It has been well-validated against invasive 'gold standard' methods (Arora et al., 2009; Baciu, Juphard, Cousin, & Bas, 2005; Benson et al., 1999; FitzGerald et al., 1997; Giussani et al., 2010; Janecek et al., 2013; Liegeois et al., 2002; Roux et al., 2003; Suarez et al., 2009), with a recent meta-analysis showing a concordance rate of 81% between fMRI and the Wada procedure (Bauer, Reitsma, Houweling, Ferrier, & Ramsey, 2014). Functional MRI measures of language lateralisation also show high test-retest reliability (Fernandez et al., 2003), and provide a better prediction of post-operative language outcome than clinical variables such as age of onset or preoperative language ability (Sabsevitz et al., 2003). Further, due to extensive work in this area, methods for good practice have been established (Seghier, 2008). Principally, threshold-independent and regional (rather than hemispheric) measures of language laterality are most reliable. The latter is especially pertinent for the epilepsy population, who may demonstrate crossed laterality in frontal and temporal regions (Thivard et al., 2005). It is also important to use large ROIs for epilepsy patients, to capture variability in location of activation (Gaillard et al., 2007; Liegeois et al., 2004; Rosenberger et al., 2009). Even investigations of typical language development have shown a greater focus on the development of lateralisation than localisation (Balsamo et al., 2006; Berl et al., 2014; Everts et al., 2009; Holland et al., 2007; Szaflarski, Holland, et al., 2006; Szaflarski, Rajagopal, et al., 2012).

3.2.1 Summary: Accurate and reliable assessment of language lateralisation?

These studies provide encouraging evidence to suggest fMRI assessments of language lateralisation are established in the pre-surgical setting, and provide an accurate and reliable alternative to invasive procedures.

3.3 Localising critical language sites

A small number of studies have attempted to validate fMRI for pre-surgical language mapping with ICM (de Ribaupierre et al., 2012; FitzGerald et al., 1997; Kim, Privitera, & Szaflarski, 2011; Lurito, Lowe, Sartorius, & Mathews, 2000; Roux et al.,

2003; Ruge et al., 1999; Rutten, Ramsey, van Rijen, Noordmans, & van Veelen, 2002). These studies suggest fMRI is good at identifying positive ICM language sites, with sensitivity ranging from 59% (in only one study; Roux et al., 2003) to nearer 100% in the majority of studies. Functional MRI results are over inclusive however, with specificity around 54%.

Methodological limitations make it difficult to draw any firm conclusions regarding the validity of fMRI for localising critical language sites in these studies. For example, the concordance of fMRI with ICM varied according to statistical threshold (FitzGerald et al., 1997; Rutten et al., 2002), the number, type and similarity of fMRI tasks used (Kim et al., 2011), and modality (auditory tasks appear more sensitive than visual) (FitzGerald et al., 1997). Further, all of the studies reviewed here have small sample sizes, and only one included paediatric patients (de Ribaupierre et al., 2012). Various fMRI tasks were used across studies, all with low-level baseline conditions (e.g. attending to scanner noise; Fitzgerald et al., 1997) which do not control for sensory-motor processing. Further, only two studies compare fMRI and ICM using corresponding language tasks (Ruge et al., 1999; Rutten et al., 2002). The majority of studies compared findings from ICM performed during a range of "ecologically valid" language tasks (e.g. reading a book, counting, spontaneously generating speech) with results from fMRI during verb generation, semantic decision, word reading or passive listening.

3.3.1 Summary: Can fMRI localise critical language sites?

Little work has been done to validate the sensitivity and specificity of fMRI in comparison to the gold standard method for localisation of critical language sites; ICM. A review of the current literature highlights methodological inconsistencies across studies and between methods, which make it difficult to draw firm conclusions. These studies do not provide consistent evidence to suggest fMRI provides an adequate non-invasive alternative to ICM for localising language. As such, the role of fMRI for informing more complex decisions regarding the degree of risk associated with removal of a surgical target in the language dominant hemisphere is currently limited. Intracranial mapping studies suggest some task design features may enhance the sensitivity and specificity of fMRI for localising

critical language sites. For example, ICM studies emphasise the importance of mapping both semantic and sensory-motor elements of language, using a variety of tasks (Corina et al., 2010; Ojemann, Ojemann, & Lettich, 2002). Tasks must also be well practiced, to reduce the identification of non-critical language regions implicated in skill learning(Ojemann et al., 2002). Intracranial mapping studies have also identified an anatomical dissociation between auditory and visual language sites in the temporal lobe, suggesting both modalities should be tested in the pre-surgical context (Hamberger, Goodman, Perrine, & Tamny, 2001; Hamberger, Habeck, Pantazatos, Williams, & Hirsch, 2014; Hamberger, McClelland, McKhann, Williams, & Goodman, 2007; Hamberger, Seidel, Goodman, Perrine, & McKhann, 2003). Future work to develop language tasks suitable for both ICM and fMRI, which engage multiple components of the language network, have higher-level baseline conditions, and investigate both auditory and visual modalities may improve the precision of fMRI and its correspondence with ICM.

4. CONCLUSIONS

4.1 Current strengths of fMRI in the pre-surgical setting

Functional MRI studies in adults have provided a comprehensive model of the mature language network, which can be used as a baseline for comparison with adult epilepsy patients. This model provides several hypotheses to test in the developmental context, which may improve the application of language fMRI in the pre-surgical setting for children with epilepsy. Language fMRI also provides a reliable and non-invasive alternative for assessing language laterality in neurosurgical candidates.

4.2 Current limitations of fMRI in the pre-surgical setting

There are two main areas for improvement of language fMRI, in order to increase its contribution to the pre-surgical decision making process. First, our understanding of how the language network develops is limited. This is mostly due to methodological constraints which prohibit the ability to distinguish between the effects of age,

performance and effort, and prohibit identification of specific structure-function mappings. Second, fMRI tasks currently in use appear suboptimal for localising critical language sites relative to ICM, and fail to engage some of the regions often compromised by surgery (e.g. the ventral and anterior temporal regions). Functional MRI also identifies regions involved in non-linguistic processes, which support task performance. This over-inclusivity - while prudent - lacks the precision required to inform more complex surgical decisions; not only whether surgery should go ahead, but also decisions regarding the extent of resection and degree of associated risk. In order to increase the contribution of fMRI to the pre-surgical decision-making process, this thesis aims to address both of these limitations.

Aims

This thesis aims to increase the contribution of non-invasive fMRI investigations of language to the neurosurgical decision-making process for children with epilepsy. The literature reviewed in the introduction to this thesis highlighted two stages of the decision-making process where the role of language with fMRI can be extended or improved. First, new fMRI tasks, which are optimised for assessing language localisation (as opposed to language lateralisation), are required. Second, a more accurate model of typical language development is required to inform predictions about the risk posed to language by surgery performed in the transitory context of development. The aims of this thesis – designed to address both of these points – are outlined below.

1. IMPROVE THE LOCALISATION OF RECEPTIVE AND EXPRESSIVE LANGUAGE NETWORKS WITH FMRI, INCLUDING SEMANTIC PROCESSING REGIONS

I aim to develop a panel of fMRI tasks which have been optimised for pre-surgical planning, and which have been optimised for use with young children and children with cognitive impairment. Design criteria for the new task panel – based on the literature reviewed in the introduction to this thesis - are outlined below:

1.1 A task battery optimised for the pre-surgical setting should:

- Provide a flexible toolkit for mapping different parts of the language system according to location of the surgical target. Different parts of the language system which require mapping include:
    - Receptive and expressive language networks
    - Auditory and visual modalities
    - Anterior and ventral temporal regions; the most common site of resection in paediatric epilepsy surgery, not currently engaged by the majority of tasks designed for children

- Provide a measure of in-scanner performance, to assess task engagement and aid interpretation
- Provide more accurate mapping of the language system than is currently possible, by:
    - Increasing the specificity of fMRI findings using high-level sensory-baseline tasks which control for sensory-motor processing and executive demands
    - Localising candidate 'core' language regions (supporting amodal semantic processing) which may be critical to language outcome
- Enable validation of fMRI findings with invasive measures, by being compatible for use with intracranial mapping

1.2 A task battery optimised for young children and children with cognitive impairment should:

- Reduce the high scan failure rates seen in young children and children with cognitive impairment
- Provide a measure of in-scanner performance, to assess task engagement and compliance
- Provide more accurate mapping of the language system than is currently possible, by:
    - Minimising executive functioning and working memory demands
    - Providing developmentally-appropriate difficulty levels which can be matched to each individual child based on their own ability, reducing identification of regions associated with effortful task performance

2. IDENTIFY AGE-RELATED CHANGES IN THE LANGUAGE NETWORK, INDEPENDENT OF TASK PERFORMANCE ACCURACY AND EFFORT

Models of typical development are important in the pre-surgical setting, to inform predictions about the risk posed to language by surgery, in the context of development. Models of typical development may also contribute to decisions regarding the optimal timing for surgery (e.g. at what point does the language system become 'adult-like'?). Currently, findings regarding age-related changes in the language network are

confounded by age-related differences in task performance accuracy and effort. There has also been a greater focus on the development of language lateralisation as opposed to language localisation. As such, the development of specific structure-function mappings is not yet known, and the development of core linguistic (i.e. semantic and syntactic) processing is not well understood.

Using the new task panel developed to address Aim 1, I aim to improve our understanding of typical language development by localising age-related changes in regions supporting semantic and syntactic processing, while providing stringent control over task performance accuracy and effort.

3. HOW THESE AIMS ARE ADDRESSED IN THIS THESIS

This thesis is composed of 4 parts:

3.1 Development

Studies performed to design and develop a new fMRI task battery (Aim 1) are presented.

3.2 Methods

Methods associated with developing and validating the new task panel (including a child-friendly pre-scan preparation protocol and movement artefact repair methods) are introduced (Aim 1).

3.3 Validation

Here I evaluate how well Aim 1 has been addressed in a series of validation studies performed in healthy children and children with drug resistant epilepsy.

3.4 Experimental investigations

Experimental investigations which a) highlight brain networks most crucial to language outcome in children with epilepsy, and b) provide novel evidence regarding typical language development (Aim 2) are presented.

4. RESEARCH QUESTIONS AND HYPOTHESES

Specific aims are outlined at the start of each chapter. Research questions and hypotheses are discussed throughout, and are summarised in the General Discussion. Practical implications and directions for future research are discussed throughout this thesis, and are summarised in the General Discussion.

PART 1: DEVELOPMENT

Chapter 1: A systematic comparison of covert and overt speech in children using fMRI

1. ABSTRACT

Overt speech is often avoided during fMRI due to motion artefacts. In this chapter I aimed to systematically compare covert and overt versions of a word generation task in terms of; a) data quality, b) lateralisation of activation and c) regional activation patterns. Thirty eight children with drug resistant epilepsy performed verb generation covertly and overtly during fMRI scanning. Measures of data quality (in-scanner movement and subjective quality ratings) and sensitivity were compared between conditions, as were measures of language lateralisation (laterality indices). Finally, brain networks supporting covert and overt speech were compared at the group level, to identify activation specific to either condition. The overt task produced higher quality scans compared to the covert task, was more sensitive for identifying activation in core language regions on an individual basis, and provided an online measure of performance crucial for improving the yield of presurgical fMRI. Laterality indices showed strong correspondence between conditions, although activation was increasingly bilateral in temporal and motor cortex during overt speech. I conclude that the use of an overt speech response mode will be advantageous for presurgical fMRI assessments of language.

2. AIMS

To enable in-scanner task performance monitoring, and to increase compatibility with intracranial mapping methods, I aim to develop a new panel of fMRI tasks which investigate language using overt speech. This chapter aims to assess the feasibility of imaging overt speech with fMRI in children, as a more ecologically valid model of expressive language than covert speech paradigms or decision-based tasks, and to allow performance monitoring during scanning. Results from this chapter will be used to guide the design of the new task panel (presented in Chapter 2).

3. INTRODUCTION

Overt speech is commonly avoided during language fMRI due to the risk of head movement artefacts (Barch et al., 1999), despite successful mapping of motor cortex during limb movements (Tiege et al., 2009). Functional MRI studies of expressive language therefore investigate covert (silent) speech or utilise button response paradigms. Imaging overt speech may be advantageous however. First, articulated speech engages more distributed regions of the brain (Shuster & Lemieux, 2005), providing a more valid model of the neural substrates supporting real-life expressive language. Second, overt responses provide an online measure of performance, essential for interpretation of findings. This is particularly important for children with cognitive impairment, who may struggle to perform the task inside the scanner. Third, overt speech may be more sensitive for identifying language cortex on an individual basis (Barch et al., 1999). Thus, overt speech tasks may be particularly appropriate for pre-surgical fMRI assessments, when accurate identification of language cortex is crucial on an individual basis, and when the margin for error is low. As movement-related artefacts can be minimized using image processing tools (Birn, Cox, & Bandettini, 2004), I aimed to assess the feasibility of imaging overt speech during continuous fMRI scanning in a sample of paediatric epilepsy patients being considered for neurosurgical intervention. This involved assessments of in-scanner movement and data quality. Further, I aimed to systematically investigate the effect of response modality on measures used in the pre-surgical decision making process: 1) patterns of neural activation during task performance, 2) sensitivity to detect activation in core language regions, and 3) measures of language dominance (laterality indices, LIs). I

hypothesised imaging overt speech with fMRI would be both feasible and advantageous for children.

4. METHODS

4.1 Participants.

Previously collected data were retrospectively analysed for 38 children with drug resistant epilepsy aged 6-17 years, who had been referred for language fMRI as part of the routine neuroimaging protocol for the Epilepsy Surgery Programme at Great Ormond Street Hospital for Children NHS Foundation Trust, between 2009 and 2011 (Table 1). Structural lesions were identified in 34 patients, including; stroke/ischemic injury (n=4), medial temporal sclerosis (n=2), focal lesions/ tumours (n=20), and other lesions including suspected Rasmussen's encephalitis (n=8).

4.2 Neuropsychological assessment.

Intelligence was assessed in 34 patients by a Clinical Paediatric Neuropsychologist as part of the Epilepsy Surgery Programme protocol at GOSH, using the Wechsler Intelligence Scale for Children (WISC IV) (Wechsler 2004). Two patients were not assessed and one patient could not be assessed with the WISC IV due to very poor verbal abilities. Handedness was determined by the Crovitz and Zener Handedness test (Crovitz & Zener 1962) where possible, or by clinical judgement where formal testing was not appropriate.

	Mean (SD)
Sex	
Female	63%
Male	37%
Handedness (N=36)	
Left	25%
Right	72%
Bilateral	3%
Intelligence (N=28)	
Verbal IQ	87 (15)
Performance IQ	89 (17)
Epilepsy variables (N=29)	
Age of onset (years)	5.5 (3.8)
Duration (years)	7.1 (4.1)
Seizure frequency (per week)	28.1 (67.6)
Seizure lateralisation with EEG (N=30)	
Left lateralised	73%
Right lateralised	3%
Bilateral distribution	23%
Structural abnormality on MRI	
Identifiable lesion on MRI	90%
Unilobar	53%
Multilobar	47%
Left hemisphere	71%
Right hemisphere	29%

Table 1: Demographic and clinical information for 38 children with drug-resistant epilepsy, referred for pre-surgical language fMRI. Mean and SD (standard deviation) for sample descriptives. Sample sizes (N) are indicated where data were not available for all participants (N=38).

4.3 Functional MRI.

MRI data were acquired on a 1.5-T Siemens Avanto System (Erlangen, Germany). Functional MRI data were acquired using a whole brain echo-planar pulse sequence (TR= 2570ms, TE= 50ms, flip angle= 90, field of view= 192x192, slice thickness= 3mm, 1 mm inter-slice gap, slices= 30, matrix size= 64x64, voxel size= 3x3x4 mm^3). A well-validated block-design verb generation (to noun) paradigm (VG) was used, as described previously (Liegeois et al., 2002). Stimuli were transmitted through MRI compatible headphones. There were ten 30 second (s) VG blocks and ten 30s baseline blocks, resulting in a total duration of 10 minutes. Each block contained eight stimuli (with an inter-stimulus interval of 3.5s). During the VG blocks participants heard common nouns and were instructed to either think of (covert condition) or say (overt condition) a single verb associated with that noun ("a doing word that goes with it"). During the baseline condition participants listened passively to amplitude-modulated noise bursts of similar duration to the word stimuli in the active condition. Children practiced the task prior to scanning. Three consecutive runs of VG were performed in a fixed order during fMRI; runs 1 and 2 were performed covertly (the routine clinical neuroimaging protocol), run 3 was performed overtly (the experimental protocol). Each run of VG used a different stimulus list, to minimise practice effects.

4.4 Image processing.

Functional images were analysed using the Statistical Parametric Mapping software (SPM8; Wellcome Department of Imaging Neuroscience, London, UK; http://www.fil.ion.ucl.ac.uk/spm). Artefacts were repaired using ArtRepair (Mazaika et al. 2005). Group spm t maps were generated using a random effects model. Three patients were excluded from group maps due to poor quality covert VG scans.

4.5 Regions-of-interest (ROIs)

I created four language regions of interest (ROIs) for the analysis of laterality indices and sensitivity, using MarsBar software (Brett et al. 2002) and the automatic anatomic labels (Tzourio-Mazoyer et al. 2002) available in the Wake Forest Pick Atlas (Maldjian et al. 2003; Maldjian et al. 2004). They included; 'Broca's area' (BA 44, 45, 47), 'extended Broca's area' (also including precentral and middle

frontal gyri), the motor cortex (precentral gyrus) and the whole temporal lobe. These regions are used clinically for investigations of language lateralisation in pre-surgical candidates.

4.6 Outcome measures

4.6.1 Performance

Behavioural performance on the verb generation task was measured a) prior to scanning, and b) inside the scanner using a sensitive head-mounted microphone with active noise cancellation (http://www.mr-confon.de/en/technology.html). Correct trials were those on which participants generated a semantically-associated verb.

4.6.2 Head movement

Measures of in-scanner movement per scan (mm/TR) were obtained using ArtRepair software (Mazaika, Whitfield, & Cooper, 2005).

4.6.3 Quality ratings

I performed subjective quality ratings for each individual t-map ($p<0.001$, uncorrected) adhering to a 4-point scale based on extent of activation within language regions (Price, 2010) and levels of noise in the data (significant voxels outside these regions). Ratings were repeated by a blind expert reviewer for a third of the dataset, and inter-rater reliability analyses were performed to ensure my ratings were accurate and reliable.

4.6.4 Sensitivity to detect activation

As a measure of sensitivity, I calculated the percentage of participants showing significant activation ($p<0.001$ uncorrected, cluster size >20) in Broca's area (inferior frontal gyrus), Wernicke's area (posterior superior temporal lobe), and the superior parietal lobe (as a non-language control region). For patients with atypical language lateralisation, activation in right hemisphere homologues was credited for this analysis.

4.6.5 Laterality indices (LIs)

Threshold independent LIs were calculated in four language regions of interest (section 4.5) using the LI toolbox, where LI=(L–R)/(L+R) (Wilke & Lidzba 2007; Wilke & V. J. Schmithorst 2006). Laterality indices >+0.2 were considered left lateralised, LIs < -0.2 right lateralised and LIs between -0.2 and 0.2 indicated bilateral representation, consistent with previous studies (Gaillard et al., 2004) . Language lateralisation was determined according to LI values in extended Broca's area during covert VG (as per the routine clinical classification of language laterality at GOSH). The typical group (left language lateralisation) included 25 patients (66%); the atypical group (right or bilateral language lateralisation) included 13 patients (34%).

4.7 Statistical analysis

Ordinal data (i.e. quality ratings) and continuous data with a non-normal distribution were compared between the three fMRI runs using Friedman's test; a non-parametric repeated-measures ANOVA. Post-hoc comparisons were performed using the Wilcoxon signed rank test. The number of participants categorised as having typical or atypical language lateralisation, and the number of participants showing activation in language regions of interest were compared between response modalities using Cochran's Q: a non-parametric test for comparing distributions for dichotomous data. Inter-rater reliability of quality ratings were assessed using Gamma; a measure of agreement for ordinal data. Gamma reflects the proportion of concordant-discordant pairs (P-Q) over the total number of pairs (P+Q), ignoring tied pairs. A positive value indicates the number of concordant pairs is larger than the number of discordant pairs (P>Q), a negative indicates the number of discordant pairs is greater than the number of concordant pairs (Q>P). Correlation analyses were performed using the non-parametric Spearman's (rho) correlation coefficient. All p-values were corrected for multiple comparisons using the Bonferroni procedure. When comparing quality ratings and LI values, overt VG was compared to covert run 2 to reduce order effects.

5. RESULTS

5.1 Effect of response mode on fMRI activation

Group maps included 36 patients; two were excluded due to poor scan quality. Covert and overt VG fMRI activations were largely overlapping in both groups (Figure 4). In the typical group both conditions produced activation in Broca's area, left precentral gyrus, basal ganglia and thalamus, bilateral posterior superior temporal lobe, insula, anterior cingulate cortex and supplementary motor area. In the atypical group both conditions produced activations in the right inferior frontal and superior temporal gyri. Condition specific activation loci are listed (Table 2).

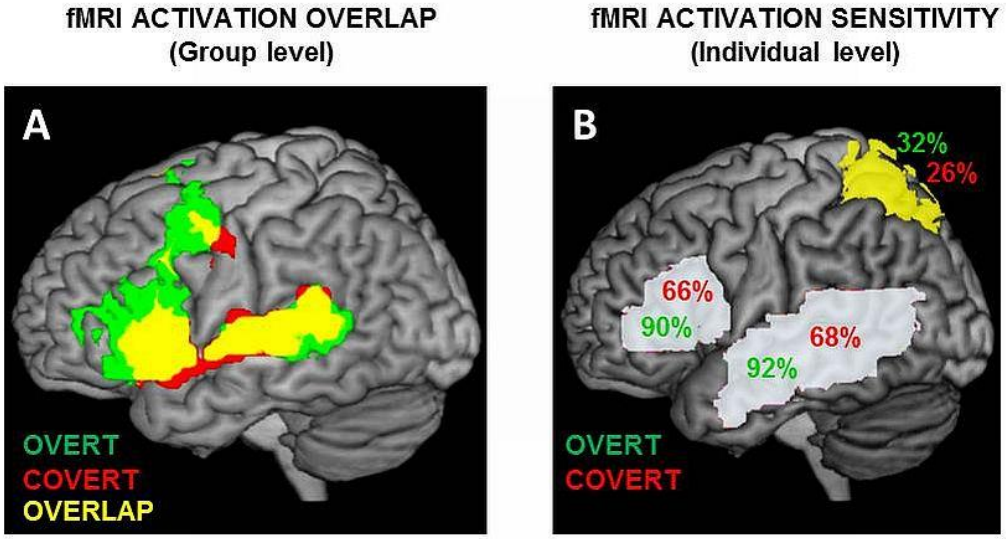


Figure 4: Effect of response modality on fMRI activation (A) and sensitivity for detecting activation in core language regions and a non-language control region (B). A: Group activation map for patients with typical language lateralisation performing overt (green) and covert (red) verb generation, showing regions active during both conditions (yellow). B: The percentage of participants showing significant activation (p<0.001, k>20) in Broca's area and Wernicke's area (white) and a non-language control region (yellow) during overt (green) and covert (red) verb generation.

	Typical language lateralisation (*N*=23)				Atypical language lateralisation (*N*=13)			
Contrast	**Cluster location**	**Co-ordinates**	**Size**	***t***	**Cluster location**	**Co-ordinates**	**Size**	***t***
Overt > covert	Left brainstem (midbrain)	-9, -7, -11	35	6.2				
	Left pre/post-central gyri	-63, -7, 22	15	4.4	Right precentral gyrus	42, -10, 37	6	4.4
	Left dACC (BA 24/32)	-12, 14, 28	13	4.3				
Covert > overt	Left SPL	-33, -55, 58	6	4	Left SPL	-15, -64, 55	20	5.2
					Right SPL	15, -58, 58	14	4.4
					Left IPL	-30, -43, 49	34	6.1

Table 2: Functional MRI activation associated with overt versus covert verb generation. Group map activation peaks for patients with typical and atypical language lateralisation, including; cluster location, coordinates in MNI space (*x*,*y*,*z*), cluster size (number of voxels) and the *t* value at a threshold of $p<0.001$ and $k>20$, uncorrected. Dorsal anterior cingulate cortex (dACC), superior parietal lobe (SPL) and inferior parietal lobe (IPL).

5.2 Outcome measures

5.2.1 Performance

Patients performed the VG task more poorly inside the scanner (54% correct) than during practice (82% correct), (p<0.001, Wilcoxon signed rank test). In-scanner performance improved with age (Spearman's rho=.49, p=0.004).

5.2.2 In-scanner movement

Movement increased across all three runs of VG ($\chi^2(2)$= 29.3, p<0.001, Friedman's test) and significantly so during overt VG (p<0.02, Wilcoxon signed rank test; Figure 5). Younger participants moved more (rho= -.58, p<0.001).

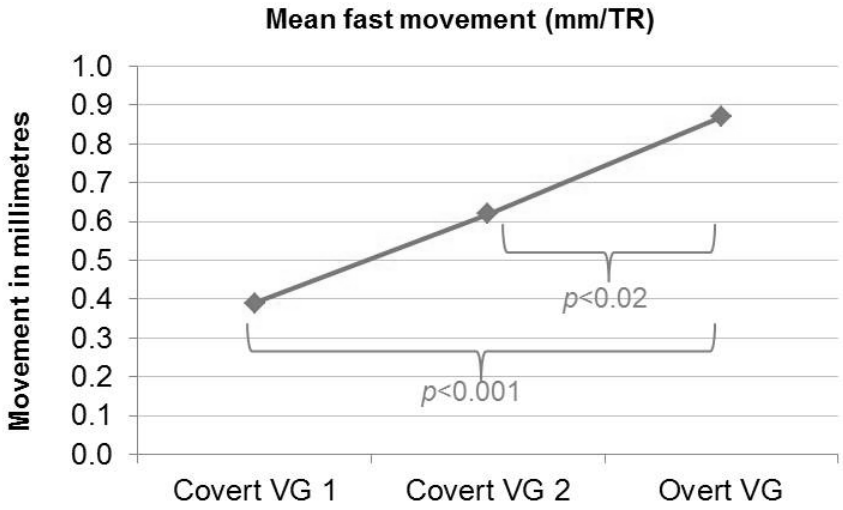


Figure 5: In-scanner movement across time, and between conditions. Participants performed two runs of covert verb generation (VG), and one run of overt VG.

5.2.3 Scan quality

Quality ratings showed high concordance between the two independent reviewers (Gamma=.86, p<0.001). Quality differed between runs ($\chi^2(2)$=6.45, p=0.04, Friedman test). *Post-hoc* comparisons showed that fewer covert scans were rated

as satisfactory/good quality (55%) compared to overt scans (76%) ($p<0.04$, Wilcoxon signed ranks test).

5.2.4 Sensitivity to detect activation

More patients showed activation in Wernicke's area (or its right hemisphere homologue for the atypical group) during overt VG (92%) compared to covert VG (68%) ($p<0.05$, Cochran's Q; Figure 4B). A similar trend was found in Broca's area (or its right hemisphere homologue for the atypical group); 90% versus 66%, respectively ($p=0.096$, Cochran's Q). There was no significant difference in sensitivity between conditions for the superior parietal control region ($p=0.56$, Cochran's Q).

5.2.5 Laterality indices

LIs were investigated in 32 patients who showed good scan quality in both conditions. There was an interaction effect of response modality and group on LI values (two-way mixed ANOVA; $F=8.54$, $p=0.008$): In the typical group LIs were reduced (less left lateralised) in the temporal lobe in the overt condition (Table 3). A similar trend was found in extended Broca's area and the motor cortex. Despite this, categorical language dominance judgements (typical versus atypical, according to LI cut-offs of >0.2 and <-0.2, respectively) were not changed by response modality ($p=0.26$, Cochran's Q). Further, absolute agreement in LI values between response modalities was high for each ROI (single measures intra-class correlation coefficients=>0.63, $p<0.001$; average measures=>0.79, $p<0.001$). On an individual basis, 3 patients (9%) showed reversed LI values for covert and overt VG in Broca's area, and 2 (6%) had reversed LIs in the temporal lobe (Figure 6).

		Mean Laterality Indices (SD)		
ROI	**Group**	**Covert**	**Overt**	***p***
Broca's area	**Typical**	0.64 (0.34)	0.39 (0.48)	-
	Atypical	-0.27 (0.55)	-0.26 (0.64)	-
Broca's area extended	**Typical**	0.71 (0.21)	0.45 (0.39)	*0.039*
	Atypical	-0.03 (0.62)	-0.32 (0.6)	-
Motor cortex	**Typical**	0.56 (0.45)	0.29 (0.56)	*0.017*
	Atypical	-0.09 (0.54)	-0.17 (0.54)	-
Temporal lobe	**Typical**	0.38 (0.4)	0.14 (0.47)	*0.002**
	Atypical	0.09 (0.54)	-0.06 (0.57)	-

Table 3: Lateralisation of language during covert and overt verb generation. Mean and standard deviation (SD) of laterality indices calculated in four language regions of interest (ROIs), for patients with typical (N=23) and atypical (N=9) language lateralisation. Significant differences between conditions are indicated with p values (post hoc two-tailed t-tests), those surviving Bonferroni correction for multiple comparisons are labelled (*). Laterality indices range from -1 to 1, where -1 indicates complete rightward lateralisation of activation, 1 indicates complete leftward lateralisation of activation, and indices between -0.2 to 0.2 indicate bilateral language activation.

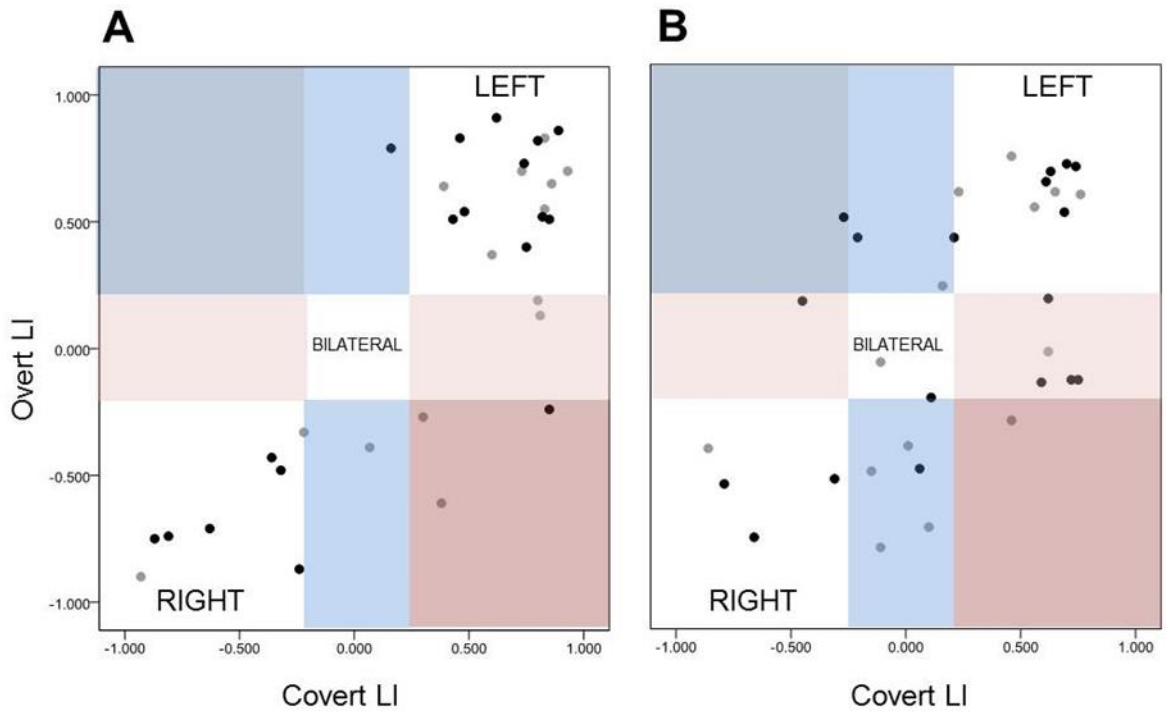


Figure 6: The effect of response modality on individual measures of language laterality in children with medically resistant epilepsy. Laterality indices (LI) from the covert and overt verb generation (VG) tasks, in Broca's area (A) and the temporal lobe (B), in 32 patients with satisfactory (grey) and good (black) quality data. Patients in the white squares have consistently lateralised (left, right or bilateral) LIs across both covert and overt conditions (Broca's area N=24, 75%; temporal lobe N=16, 50%). Patients within the four light coloured strips (horizontal red and vertical blue) also have consistent LI values between conditions; in the bilateral range. For patients in the two darkest squares (bottom right red square and top left blue square) LIs are reversed between conditions (Broca's area N=3, 9%; temporal lobe N=2, 6%). Note the majority of these patients have satisfactory data quality, rendering LI measures less reliable relative to those with good data quality.

6. DISCUSSION

In this chapter, I demonstrate the feasibility of imaging overt speech with continuous fMRI scanning in children with epilepsy, and provide an initial step towards validating overt language tasks for pre-surgical language fMRI assessments. Through a systematic comparison of covert and overt language tasks, I also show it is advantageous to

investigate overt language in the pre-surgical setting; the overt task produced a larger proportion of satisfactory/good quality scans, and proved more sensitive for identifying language regions on an individual basis. These findings are even more persuasive when one considers overt data were acquired last in the task protocol, when temporal order effects (such as increasing in-scanner movement and decreasing concentration) may reduce scan quality. Further, the in-scanner performance monitoring it provided was valuable, highlighting the contribution of poor task performance to scan failure.

6.1 The feasibility of imaging overt speech fMRI.

Previous studies have used sparse fMRI sampling techniques in an attempt to avoid movement artefact (Vannest et al., 2010). This method is sub-optimal for the clinical setting however (due to longer scanning time) and also appears unnecessary: The increase in head movement associated with overt speech (0.25mm on average) during continuous scanning did not compromise data quality. Furthermore, when excessive movement (>2mm) occurred (N=4) it was more likely to be seen throughout the entire scanning period (N=3) rather than being restricted to the overt condition (N=1). Crucially, there was a 76% success rate for imaging overt speech in this cohort of children with significant cognitive impairment and behavioural co-morbidities. A covert version of the same task produced a significantly lower success rate; 55%. As such, not only is it feasible to investigate overt speech with continuous scanning, it appears advantageous for improving the yield of fMRI.

6.2 The advantages of imaging overt speech with fMRI.

6.2.1 Online performance monitoring.

A previous large scale study of fMRI success rates in children with epilepsy suggests 16% of children fail at least one run of language fMRI (Yerys et al., 2009). While the authors suggest 39% of these scan failures are due to excessive in-scanner movement, they cannot account for the remaining 61%. Scan failure rates in this chapter closely correspond to the findings of Yerys and colleagues; 16% of children had failed language fMRI results across the entire battery of tasks performed. Also consistent with Yerys, 33% of scan failures reported in this

chapter were due to excessive in-scanner movement. Due to the online performance monitoring afforded by the overt speech response mode in this study, I was able to show the remaining proportion of scan failures were associated with poor in-scanner performance (<58% accuracy). Thus, I provide evidence to suggest poor in-scanner performance may be the cause of up to 61% of scan failures in children with drug-resistant epilepsy. Because language tasks with an overt speech response mode provide a mechanism by which performance difficulties can be identified and rectified 'online', they have the potential to prevent the majority of scan failures in children with epilepsy. A button-press response mode can be used to monitor performance as an alternative to overt speech. However, a forced-choice button-press response mode may inflate estimates of performance accuracy. Further, button-press response tasks usually require a decision-making element, which may engage additional (non-linguistic) components to support task performance. The use of an overt speech response mode therefore allows performance monitoring during a purely linguistic task. My findings also reiterate the importance of having the opportunity for extensive practice on language tasks prior to scanning for patients with cognitive impairment. This is especially true for younger patients, who struggle most. Mock scanning procedures which are used for very young children (Raschle et al., 2009) may also be beneficial for children with cognitive impairment, and may help improve scan success rates (de Bie et al., 2010). The role of mock scanning procedures in paediatric fMRI are discussed further in Chapter 3, where I also introduce a new pre-scan preparation protocol for use with the Panda Games task battery.

6.2.2 A more ecologically valid model of expressive language.

Overt speech paradigms may offer a more ecologically valid model for studying expressive language, beneficial for the pre-surgical setting. However, there is debate whether additional brain regions engaged during overt speech (relative to covert speech) reflect higher order language processing (Kielar, Milman, Bonakdarpour, & Thompson, 2011; Shuster & Lemieux, 2005), or motor aspects of speech production (Vannest et al., 2010). In support of the findings from Vannest et al., (2010), I found regions more active during overt speech were

related to the motoric act of speech (Figure 7). These included; a) regions of the dominant precentral gyrus associated with movements of the lower face and larynx (Grabski et al., 2011), b) the dorsal aspect of the left anterior cingulate cortex (BA 24 and 32), known as the vocalisation region (Paus, 2001), and c) the left midbrain, in proximity to the medial geniculate nucleus, involved in phonemic processing of speech (von Kriegstein, Patterson, & Griffiths, 2008). These findings suggest that while covert speech paradigms may provide a sufficient model of higher-order processing during single-word retrieval, overt speech paradigms may be beneficial in the pre-surgical decision-making process by providing a more comprehensive map of the wider expressive language network; identifying brain regions involved in both word selection *and* production processes. However, whether this translates to improved post-surgical outcome is yet to be seen.

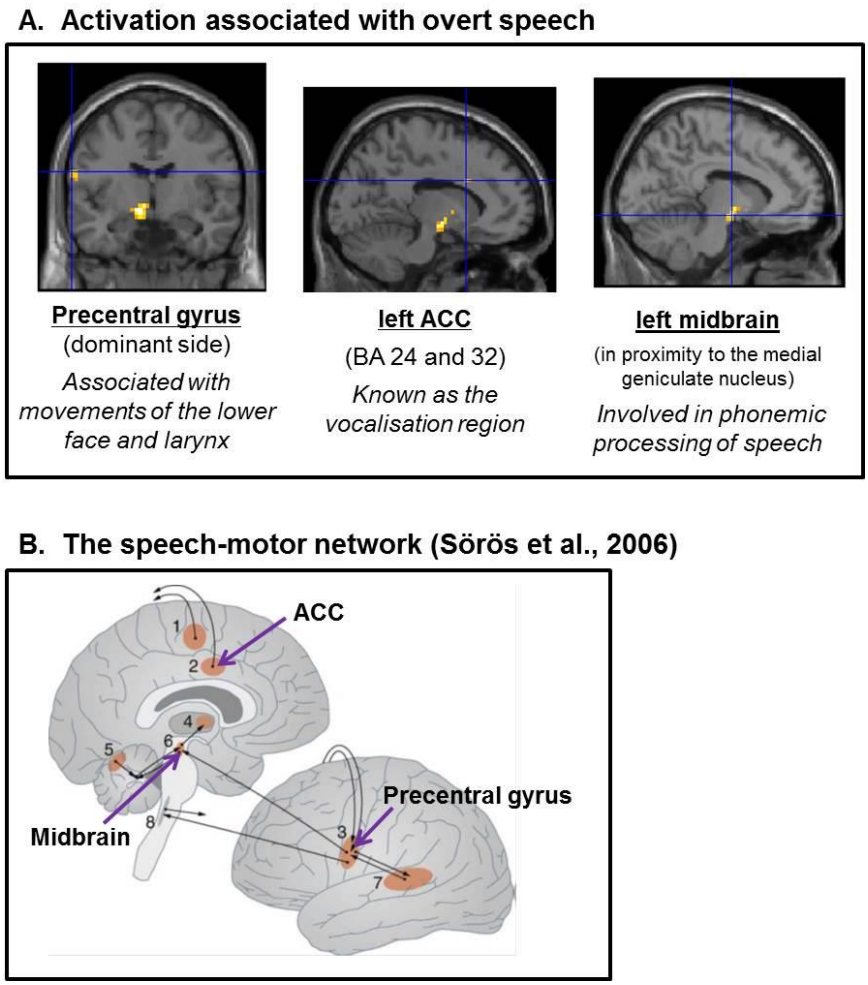


Figure 7: Activation associated with overt speech (A), in comparison to the speech-motor network (B). Results of group map comparisons in the typical lateralisation group (N=32) showing brain regions significantly more active during the overt condition relative to the covert condition are shown (A), at $p<0.01$ uncorrected. Regions of overlap between my findings reported here and regions of the speech-motor network (Soros et al., 2006) are indicated by purple arrows (B).

6.3 Validation of overt speech tasks for language lateralisation in the clinical context.

As a first step towards validation for use in the pre-surgical setting, I compared LIs from the overt task to LIs from the covert task; which has been validated for pre-

surgical language fMRI using invasive measures (Arora et al., 2009). Overt VG was associated with more bilateral involvement of motor and temporal cortices, which may reflect articulation and processing of auditory feedback under challenging conditions (Tourville, Reilly, & Guenther, 2008), respectively. Encouragingly, despite the subtle increase in bilateral representation, the overt task was successful in determining individual language dominance when compared to the covert task. These findings require validation with invasive language mapping and post-operative outcome data (as discussed in more length in Chapter 6). I hypothesise the overt response mode will increase the correspondence of fMRI with invasive measures which also probe overt speech, although this requires empirical support.

7. CONCLUSIONS

The increased yield and sensitivity of overt language fMRI relative to covert language fMRI in this representative sample of children with drug resistant epilepsy suggests language tasks which use an overt response mode should improve the diagnostic workup of patients prior to surgery. It will therefore be both feasible and advantageous to include an overt speech-response mode in the new Panda Games task battery.

Chapter 2: Design, development and pilot of a new fMRI task battery called 'The Panda Games'

1. ABSTRACT

In order to investigate developmental changes in the language network using fMRI, it is important to control for age-related differences in performance accuracy and difficulty (effort). In this chapter I introduce a new series of language tasks for fMRI called the Panda Games. I aimed to collect performance, age of acquisition and word frequency data for the new series of stimuli used in the Panda Games, and to use this information to create several age-appropriate difficulty levels for each task. Healthy children (N=29, 5-14 years) and adults (N=9, 23-55 years) performed the Panda Game outside the scanner. Reaction time, response duration and accuracy data were captured. Name agreement, the H statistic (a measure of spread in responses), age of acquisition and word frequency were calculated. I then grouped stimuli using established predictors of performance accuracy and difficulty from the naming literature (word frequency, name agreement, H, and age of acquisition data). This produced four difficulty levels (very easy, easy, moderate and difficult) for each of the three tasks piloted. Between-level comparisons showed children performed each level with increasing difficulty (longer reaction times) and with less accuracy. Encouragingly, difficulty levels were well matched for variables known to affect fMRI results; stimulus length, response length, and the proportion of living/non-living and natural/manmade items. I conclude that task performance is poorer in children compared to adults, and that age-appropriate task design should enhance the yield of presurgical fMRI in children with epilepsy and cognitive impairment.

2. AIMS

In this chapter I present a proposal for a new fMRI task battery (The Panda Games), optimised for pre-surgical investigations of language in children with epilepsy. I also present findings from a behavioural pilot study of 140 new picture and sentence stimuli, for use in this task battery. To make the Panda Games suitable for children aged 5-16 years, this chapter had four main aims:

1. Collect performance, age of acquisition and word frequency data for a series of new, child-friendly stimuli in children aged 5-16 years, as per previous naming norming studies.
2. Investigate the necessity of using developmentally-appropriate tasks to investigate language, by comparing performance between children and adults.
3. Create four difficulty levels within each of the Panda Games; very easy, easy, moderate, and difficult.
4. Establish which grouping variable – age or verbal ability – best explains task performance in children, and can be used to assign children to a given difficulty level for each of the Panda Games.

3. INTRODUCTION

In order to investigate age-related changes in the language network, it is important to control for performance and effort (Gaillard, Grandin, et al., 2001). Despite this, many language fMRI studies adopt a one-size-fits all approach to task design, even when comparing adults and children (Brauer & Friederici, 2007; Holland et al., 2007; Wilke et al., 2006). These studies typically report age-related changes in language network activation consistent with decreasing bilateral activity and increasing activation in the language dominant hemisphere, particularly frontal regions. However, it is difficult to establish whether these changes are associated with age specifically, or reflect differences in effort or performance accuracy. Studies which have used developmentally-appropriate language tasks are scarce, but suggest age-related changes in the core language network may be more subtle than previously expected after 4 years of age (Berl et al., 2010; Berl et al., 2014). For pre-surgical mapping (and for experimental purposes) it is of interest to identify changes in the core language network

specifically associated with maturation; allowing clinicians and scientists to make accurate and reliable predictions from neuroimaging in the context of on-going development.

3.1 Measures of performance accuracy and performance difficulty (effort).

Performance accuracy (i.e. error rates) can be measured directly, and can therefore be controlled. In contrast, the effort required to perform a task is more difficult to quantify. Reaction time however has been used as a proxy for performance difficulty – or effort - since the 19th century (Cattell, 1886); at least for word retrieval tasks. Importantly, work in the 1960s began to identify variables which influenced both performance accuracy and reaction time (effort):

3.1.1 Name agreement.

Name agreement (NA) reflects the proportion of participants that produced a given response, where a score of 1 indicates all participants produced the given response, and a score of 0 indicates no participants produced the given response. NA therefore provides a measure of how dominant a response is. This is predictive of both naming latency (Barry, Morrison, & Ellis; Snodgrass & Yuditsky, 1996) and naming accuracy (Adlington, Laws, & Gale, 2009).

3.1.2 H statistic.

First introduced as the *U* statistic (Gilhooly & Logie, 1980; Lachman, Shaffer, & Hennrikus, 1974), *H* is similar to (and highly correlated with) NA, and is a measure of response uncertainty. Response uncertainty refers to the concept that while a written word evokes a single spoken name, a picture may produce several candidate names (Fraisse, 1969). Picture naming is therefore associated with a degree of uncertainty. Relative to picture naming, naming based on an item description (as in our RN and LN tasks) may produce even more candidate names and therefore greater response uncertainty. This degree of uncertainty can be measured as the number and frequency distribution of alternative names for a given picture. The *H* statistic has been found to be an important predictor of naming latency when it is >2 (Lachman, 1973; Lachman et al., 1974). In a more recent and large scale study, the *H* statistic has also been found to correlate

strongly (-.83) with picture naming performance accuracy (Adlington et al., 2009).

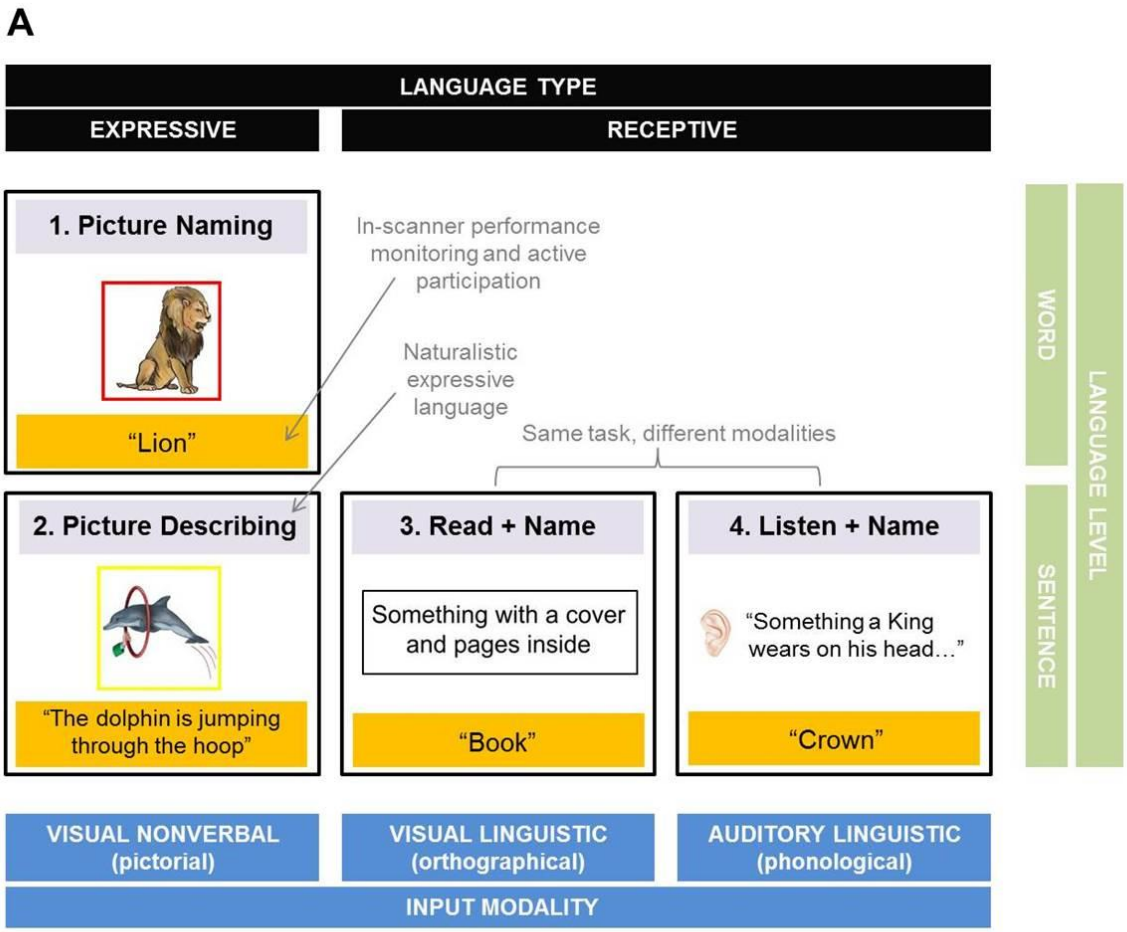
3.1.3 Word frequency.

Word frequency (freq) - printed or spoken - has consistently been shown to relate to picture naming latency, such that higher frequency words are more easily retrieved than lower frequency words (Adlington et al., 2009; Barry et al.; Oldfield & Wingfield, 1965; Snodgrass & Yuditsky, 1996). This is also holds true for reading (Frederiksen & Kroll, 1976; Huttenlocher & Kubicek, 1983).

3.1.4 Age of acquisition.

Ratings of the age at which a name was acquired (age of acquisition; AoA) is another crucial predictor of naming latency (Adlington et al., 2009; Barry et al.; Carroll & White, 1973; Holmes & Ellis, 2007; Johnston, Dent, Humphreys, & Barry, 2010; Snodgrass & Yuditsky, 1996), as well as naming errors (Holmes & Ellis, 2007) and name agreement (Snodgrass & Vanderwart, 1980). Some authors argue AoA may even be a stronger predictor of naming latency than word frequency (Carroll & White, 1973). Further, in the only study to investigate picture naming in children as young as 5 years, AoA (along with NA) was the only predictor of naming accuracy (Cycowicz, Friedman, Rothstein, & Snodgrass, 1997). Although these AoA ratings are subjective, they correlate strongly with objective measures of AoA (r values between .85 and .93) supporting their validity (Carroll & White, 1973; Gilhooly & Logie, 1980).

I have designed a new fMRI task panel ('The Panda Games', Figure 8) for investigating language, which has been optimised for children and pre-surgical mapping according to recommendations from the literature reviewed in the introduction to this thesis. In this chapter I aim to pilot stimuli for the Panda Games, and create subsets of stimuli (difficulty levels) appropriate for children of different ages according to established predictors of performance accuracy and difficulty from the large body of naming literature (discussed above). In doing so, I aim to create age-adjusted tasks which control for some of the between-subject variance in performance accuracy and effort when investigating language network development.


A
LANGUAGE TYPE
EXPRESSIVE
RECEPTIVE
1. Picture Naming
"Lion"
In-scanner performance monitoring and active participation
Naturalistic expressive language
Same task, different modalities
2. Picture Describing
"The dolphin is jumping through the hoop"
3. Read + Name
Something with a cover and pages inside
"Book"
4. Listen + Name
"Something a King wears on his head…"
"Crown"
WORD
SENTENCE
LANGUAGE LEVEL
VISUAL NONVERBAL (pictorial)
VISUAL LINGUISTIC (orthographical)
AUDITORY LINGUISTIC (phonological)
INPUT MODALITY

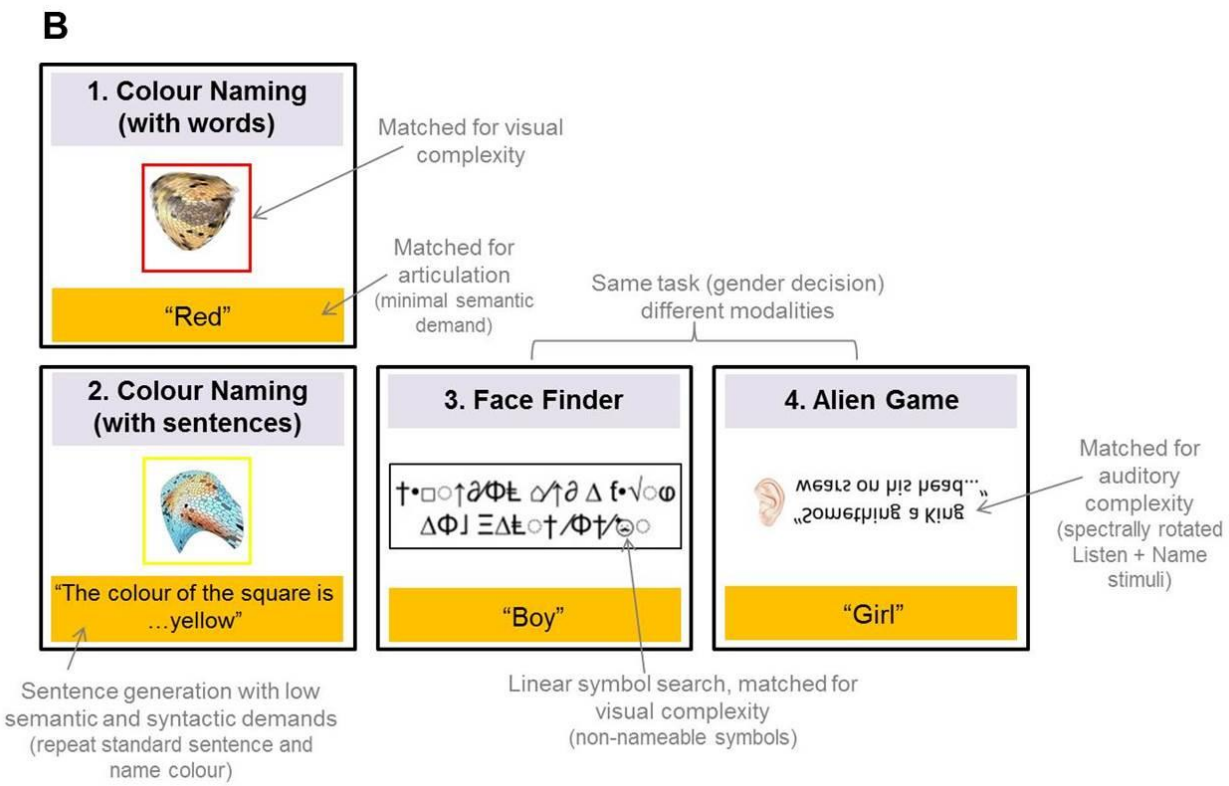

B
1. Colour Naming (with words)
"Red"
Matched for visual complexity
Matched for articulation (minimal semantic demand)
Same task (gender decision) different modalities
2. Colour Naming (with sentences)
"The colour of the square is …yellow"
3. Face Finder
"Boy"
4. Alien Game
"Girl"
Matched for auditory complexity (spectrally rotated Listen + Name stimuli)
Sentence generation with low semantic and syntactic demands (repeat standard sentence and name colour)
Linear symbol search, matched for visual complexity (non-nameable symbols)

Figure 8: (previous page) The Panda Games: Proposal of a new fMRI task battery for investigating language networks in children. Language tasks (A) and sensory-motor baseline tasks (B) are presented, including child-friendly names (purple boxes), example stimuli, and the expected participant response (yellow boxes). Notes regarding specific design features are in grey.

A brief description of The Panda Games (also see Chapter 3, section 9): During *Picture Naming* (PN) participants view pictures of animals and objects, and say the name. During *Picture Describing* (PD) participants view pictures of events, and describe what is happening in one simple sentence (subject-verb-object structure). During *Read and Name* (RN) participants silently read simple sentence descriptions of animals and objects, and say the name of the item being described. During *Listen and Name* (LN) participants hear simple sentence descriptions of animals and objects, and say the name of the item being described. During *Colour Naming with words* (CNw) participants name the colour of the square border. During *Colour Naming with Sentences* (CNs) participants use the same sentence structure ("The colour of the square is…") to name the colour of the square border. During *Face Finder* (FF) participants search along a string of non-nameable symbols to find the smiley face, and say whether it is a boy or girl. During the *Alien Game* (AG) participants listen to 'aliens' talking (spectrally rotated versions of LN stimuli), and say if the alien is a boy or girl.

4. METHODS

4.1 Participants

A total of 29 children (5-14 years, mean = 8.86, SD=2.48, 66% female, 76% right handed, Mean VIQ = 105, SD=18.87) and 9 adults (23-55 years, mean=36.38, SD=14.18, 67% female, 78% right handed) were recruited through general advertisement and from primary schools. All participants spoke English as their first language, had normal or corrected-to-normal vision and hearing, and had no history of neurological or cognitive impairment or learning disability. All adults and ten children performed all three tasks; Picture Naming (PN), Picture Describing (PD) and Read and Name (RN). To reduce testing time, 19 children performed either a) PN and PD, or b) RN. Order of task performance was randomised. For children under seven years of age, sentences in the RN game were read aloud by the investigator. In total, 20 children performed each task. These were matched for age, VIQ, sex and handedness across tasks:

4.1.1 Picture Naming

20 children aged 5-14 years (mean=9.35, SD=2.56), mean VIQ=108 (SD=16.31), 70% female, 80% right handed.

4.1.2 Picture Describing

20 children aged 5-14 years (mean=9.35, SD=2.56), mean VIQ=108 (SD=16.31), 70% female, 80% right handed.

4.1.3 Read (or listen) and Name

20 children aged 5-14 years (mean=9.2, SD=2.76), mean VIQ=106 (SD=19.08), 65% female, 70% right handed.

4.2 Language tasks

4.2.1 Picture Naming (PN)

Participants were asked to name 140 pictures, depicting objects and animals (Figure 9). Picture stimuli were developed by Prof Cathy Price and colleagues, for neuroimaging investigations of language in healthy and patient populations

(see Appendix 1.1). Participants were asked to name items as quickly as they could, and to say the first thing that "popped into their head" (for task instructions see Appendix 2.).

4.2.2 Picture Describing (PD)

Stimuli from the Picture Naming task were paired by Price and colleagues to create 70 event images (Figure 9). Each stimulus showed a subject and an object, with one of four actions; eating, drinking, jumping or falling (see Appendix 1.3). Participants were asked to describe what was happening in the picture in one simple sentence, using a subject-verb-object (S-V-O) structure. Participants were asked to respond in the active present tense ("as if it's happening now"), e.g. "The squirrel *is jumping* from the tree", as opposed to "The squirrel *jumped* from the tree". For drinking stimuli, participants were asked to say what the subject was drinking from, as opposed to what the subject was drinking (for task instructions see Appendix 2.2).

4.2.3 Reading (or listening) and Naming (RN)

Single sentence descriptions were formulated for all 140 items in the Picture Naming task (Figure 9; Appendix 1.5). Sentences were taken from one of three versions of the children's Oxford English Dictionary; Oxford First Dictionary (5-7 years), Oxford Primary Dictionary (7-11 years), or the Oxford Dictionary For Schools (11-16 years). Participants were asked to read (or listen to) the sentence and name the item being described as quickly as possible (Appendix 2.3 for task instructions).

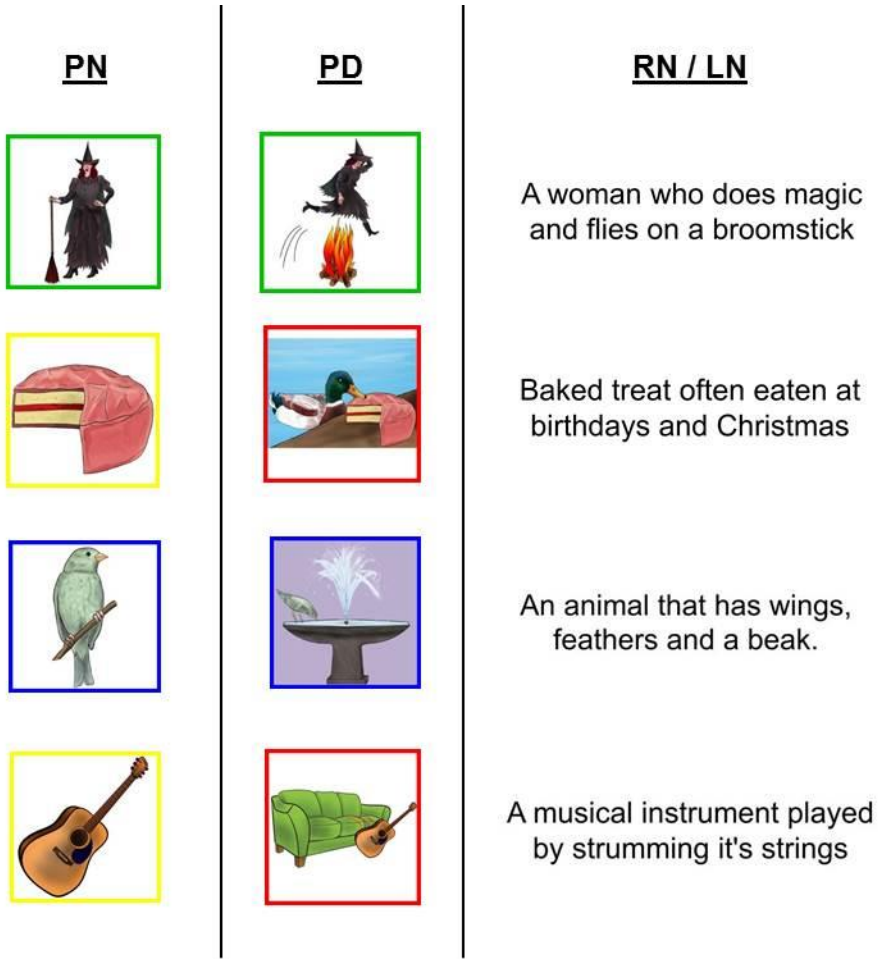


Figure 9: Stimuli examples for three language tasks. Including: Picture Naming (PN), Picture Describing (PD) and Reading/Listening and Naming (RN/LN).

4.3 Testing procedures

Stimuli were presented on a laptop using DMDX software (Forster & Forster, 2010). The first five stimuli were introduced as a practice session. If participants did not perform the task correctly during this practice session, task instructions were repeated and the task was re-started. Stimuli were presented sequentially and in a randomised order. For each trial, a fixation cross was presented in the centre of the screen (1000ms) followed by the stimulus, which remained on screen until the participant made a response. Participants responded vocally for each task. The order in which tasks were performed was counterbalanced between participants.

4.4 Neuropsychological assessment

In order to investigate whether children should be assigned to difficulty levels according to their age or verbal ability, verbal IQ was assessed in children using the Wechsler Abbreviated Scale of Intelligence (Wechsler, 1999).

4.5 Performance data and stimuli characteristics: Naming tasks (PN and RN)

4.5.1 Reaction time

For PN and PD, vocal responses and onset times were recorded using a microphone headset and voice trigger system. For RN vocal responses were recorded manually and reaction time data was captured using a button-press response (the voice trigger was not used due to a high false trigger rate for this task, associated with participants saying 'umm' or moving their lips during silent reading). False triggers of the microphone (RT <100ms), failed triggers (as noted during testing or indicated by RT>3500ms), missed responses (MISS) and 'do not know' (DNK) responses were excluded from this analysis.

4.5.2 Dominant response

The dominant response represents the modal response produced in the child group, and will be taken as the 'correct' answer for future neuroimaging studies using these stimuli in developmental populations.

4.5.3 Errors

MISS, DNK responses and incorrect names (i.e. names which were not semantically related to the picture) were classed as errors.

4.5.4 Word frequency (freq)

Word frequency data was obtained for the dominant response for each stimulus, from the British National Corpus ("The British National Corpus," 2007)(BNC XML Edition, 2007: http://www.natcorp.ox.ac.uk/); a 100 million word collection of samples of written and spoken language from a broad range of sources,

designed to represent a wide cross-section of British English from the later part of the 20th century.

4.5.5 Name agreement (NA)

Name agreement reflects the proportion of participants in the child group who responded with the modal name (after excluding MISS or DNK responses), and provides a measure of how dominant the modal name was among all names produced.

4.5.6 H statistic

The *H* statistic is sensitive to how widely distributed responses are, over all of the unique names produced. An increasing *H* value indicates increased variability and frequency of names given (increased response uncertainty). *H* was calculated according to the following formula:

$$H = \sum_{i=1}^{k} p_i \log_2 \left(\frac{1}{p_i} \right)$$

Where *k* is the number of different names produced to a picture, and *Pi* is the proportion of subjects producing the name. All participants produced the same name when *H*=0.

4.5.7 Age of acquisition (AoA)

Age of acquisition data was obtained for 130 (93%) of stimuli from three British adult naming norming studies (K. Gilhooly & Logie, 1980; Johnston et al., 2010; Stadthagen-Gonzalez & Davis, 2006). In all three studies AoA data was collected using a 7-point likert scale developed by Gilhooly and Logie (1980). This scale comprised seven age bands, beginning with 0–2 years of age and increasing at 2 year intervals up to 13 years and over. For Johnston et al., (2010) participants were asked to rate the age at which they thought they had first learned the name of the object. They were instructed to rate the name and not the picture. Participants were instructed to leave this rating blank if they had failed to name the item. For

Stadthagen-Gonzales & Davis (2006), participants were asked to estimate (in years) when they had learned the word. Responses were then converted to the Gilhooly and Logie (1980) scale.

4.6 Performance data and stimuli characteristics: Picture Describing (PD)

Reaction time (RT) was calculated for PD stimuli as described for PN and RN (section 4.5.1). Measures which differed for the PD task are explained below.

4.6.1 Dominant response

For PD, the dominant response represents the modal subject-verb-object (S-V-O) structure used to describe the image (e.g. '*squirrel – jumping – tree*'). Differences in prepositions were therefore disregarded (e.g. '*The squirrel is jumping off the tree'* and '*The squirrel is jumping from the tree*').

4.6.2 Errors

Errors were defined according to the following criteria; a) failure to use a S-V-O structure, b) failure to use the active-present tense, c) responses describing what the subject was drinking (e.g. '*The man is drinking water*'), rather than what the subject was drinking from (e.g. '*The man is drinking from a cup*'), d) responses including semantically unrelated (incorrect) subject or object names, and e) DNK responses.

4.6.3 Response duration

The time taken to generate a sentence response was calculated. Responses with a duration >10,000ms or <1000ms were excluded from this analysis (as failed or false triggers, respectively) as were items with a MISS or DNK response.

4.6.4 Stimuli characteristics

Response frequency (freq), response agreement (RA) and the *H* statistic were calculated by averaging values from the dominant subject and object names.

5. RESULTS

5.1 Performance data and stimuli characteristics

The dominant name, RT, NA, *H* and AoA data for all piloted stimuli are provided in Appendix 3. Encouragingly, my results correlate significantly with those reported in a previous study of picture naming in young (5-7 years) American children (Cycowicz et al., 1997); NA (r=.31, $p<0.001$) and *H* (r=.25, $p=0.003$). Further validation cannot be performed due to the scarcity of developmental data in the picture naming field.

5.2 Task performance in children and adults

Reaction times differed between tasks (*p* values <0.01) and were longest for RN in both child and adult groups. Children had significantly longer RTs compared to adults for both naming tasks (PN and RN), and took longer to produce a sentence in the PD task (Table 4). Children produced more errors than adults on all tasks (Table 4). Error rates for adults were very low. For children, error rates were variable between tasks and were highest in the RN task.

Task performance			
	Child (N=20)	**Adult (N=9)**	***p***
Reaction time (*s*)			
PN	1.11 (19.8)	0.93 (0.14)	0.02
RN	5.23 (1.88)	2.39 (0.61)	<0.001
PD	1.56 (5.47)	1.36 (0.24)	0.35
PD (duration)	*6.58 (1.02)*	*4.90 (2.03)*	0.007
Error rate			
PN	1% (2%)	0% (0%)	0.005
RN	18% (17%)	1% (1%)	<0.001
PD	9 (8%)	2% (2%)	0.002

Table 4: Task performance data. Mean (and standard deviation) of reaction time (seconds) and error rates in children and adults during performance of three language tasks; Picture Naming (PN), Read/Listen and Name (RN) and Picture Describing (PD).

5.3 Predictors of task performance in children

Stepwise linear regression was performed to investigate which variable (age or VIQ) predicted the most variance in performance for PN and RN (as measured by RT) and for PD (as measured by error rates) in children. Age was the only significant predictor of performance in the final model for each task; PN (R_2 adjusted =.34, $F(1,17)=10.31$, $p=0.005$), RN (R_2 adjusted =.50, $F(1,16)=18.06$, $p=0.001$), and PD (R_2 adjusted =.47, $F(1,17)=16.74$, $p=0.001$). As such, I explored age further as the most valid variable for grouping children according to task performance.

5.4 Age groupings which explain task performance

I compared three age grouping variables using stepwise linear regression, to identify which grouping method explained the most variability in task performance; 1) two groups (5-10 and 11-16 years), 2) three groups (5-8, 9-12 and 13-16 years), and 3) four groups (5-7, 8-10, 11-13, 14-16 years). The four group solution was the only variable retained in the final model for PN (R_2 adjusted = .36, $F(1,18)=11.45$, $p=0.003$), RN (R_2 adjusted =.47, $F(1, 18)=17.81$, $p=0.001$), and PD (R_2 adjusted=.40, $F(1,18)=13.45$, $p=0.002$). As such, tasks in the Panda Games battery will be designed such that they are composed of four difficulty levels; one for each of these age groups. When conducting neuroimaging investigations of language in this thesis, children with verbal abilities in the normal range will be assigned to the appropriate difficulty level of each task according to these age groups (i.e. 5-7 year olds to level 1, 8-10 year olds to level 2, 11-13 year olds to level 3 and 14-16 year olds to level 4). Children whose verbal abilities fall outside the normal range (i.e. children with drug-resistant epilepsy) will be assigned to the appropriate level according to their age-equivalent level of functioning (as determined by performance on neuropsychological assessments).

5.5 Selecting stimuli to create four difficulty levels within each task

To create four difficulties within each language task (PN, RN and PD), I categorised stimuli in each task as either very easy, easy, moderate or difficult according to

variables known to effect performance difficulty from the naming literature (reviewed in the introduction of this chapter).

5.5.1 Exclusions

First I excluded stimuli which were inappropriate for children aged between 5 and 16 years (according to criteria used in previous studies from the naming literature). I excluded 19 stimuli from the PN task; 16 due to NA <60% (Barry et al 2007), and 3 due to an error rate >10% in children (as in Adlington et al., 2009 and similar to Hamberger et al., 2003). Exclusion criteria were more lenient for the RN and PD tasks as RTs and error rates were higher and more variable for these tasks compared to the classic picture naming paradigm. In total, 27 stimuli were excluded from the RN task; 14 due to a high error rate (>20% in children aged 8 years and over), nine due to low NA (<50%), and four due to both a high error rate and low NA. Nineteen stimuli from the PD task met exclusion criteria; 7 due to a high error rate (>30%), 12 due to low NA (<50%) and 4 due to both high error rates and low NA. Due to the limited number of stimuli available in the PD task however - and to try to reduce stimuli overlap between difficulty levels - these stimuli were not excluded. Instead, they were automatically assigned to the most difficult level (4).

5.5.2 Stimuli difficulty ratings

For the remaining stimuli, I calculated the interquartile range for freq, NA and *H* for each task based on weighted average percentiles (which take the median as the 50^{th} percentile) (Table 5). From here on these variables are collectively referred to as *level determinants*. For each level determinant, stimuli within each inter-quartile range were given a rating of 1 (very easy), 2 (easy), 3 (moderate) or 4 (difficult). For the *H* statistic for example, stimuli in the 1st quartile were given a rating of 1 (very easy), whereas those in the 4th quartile were given a rating of 4 (difficult). This rating system was reversed for freq and NA, such that stimuli with the highest frequency and highest NA (in the 4th quartile) were given ratings of 1 (very easy) and stimuli with the lowest freq and NA (in the 1st quartile) were given ratings of 4 (difficult). Each stimulus was therefore assigned a difficulty rating three times; one for each level determinant. The three ratings were then

averaged, to produce a mean difficulty rating for each stimulus. Because age-related differences in task performance were apparent in the error rate for the PD task, rather than in RTs (Table 4), ratings were also performed based on the error rate for this task.

	Weighted-average percentiles for level determinants					
Task	**Mean**	**SD**	**Inter-quartile range**			
			1	**2**	**3**	**4**
Picture Naming (121 stimuli)						
Word frequency	4439	8340	517	1231	3865	58821
Name agreement	0.86	0.17	0.85	0.95	1.00	1.00
H statistic	4.99	5.77	0.00	3.47	4.93	20.65
Read and Name (113 stimuli)						
Word frequency	3947	8491	517	1126	2886	58821
Name agreement	0.81	0.16	0.68	0.84	0.95	1.00
H statistic	9.44	6.79	4.40	8.51	13.80	28.53
Picture Describing (51 stimuli)						
Response frequency	3856	5689	647	1756	4544	32044
Response agreement	0.82	0.14	0.74	0.85	0.94	1.00
H statistic	8.03	5.32	4.01	7.88	11.65	19.98
Errors (%)	8%	10%	0%	5%	15%	30%

Table 5: Level determinants and their inter-quartile range across three language tasks. The number of stimuli available for each task (following exclusions discussed in section 5.5.1) is included in brackets.

5.5.3 Assigning stimuli to one of four difficulty levels

For each task, stimuli were ranked according to their mean difficulty rating (to two decimal places); the first 24 stimuli were assigned to level 1, the next 24 stimuli were assigned to level 2, and so on. Using this method I could divide each task into four (24-item) difficulty levels, for use with four age-groupings; 5-7 years (level 1), 8-10 years (level 2), 11-13 years (level 3) and 14-16 years (level 4).

5.5.4 Age of Acquisition

Finally, I ensured that all level 1 stimuli were acquired by six years of age (AoA ratings of 1 or 2), all level 2 stimuli were acquired by eight years (AoA rating 3-4 or less), all level 3 stimuli were acquired by 12 years (AoA ratings of 5-6 or less), and all level 4 stimuli were acquired by 13 years of age (AoA ratings of 7 or less).

5.6 Four language tasks with four difficulty levels: Descriptives and validity

Between-level comparisons were performed for two reasons. First, to ensure levels were matched for stimulus/response length, category (living/non-living, natural/man-made) and - for PD - action (verb). Second, to establish how much levels differed, according to measures of task difficulty; a) level determinants, b) RTs, and c) error rates. There was a 45% overlap of stimuli between levels in the PD task. The number of unique items in levels 1-4 were; 20 (83%), 8 (33%), 2 (8%) and 14 (58%), respectively. For the purpose of comparing stimuli characteristics between levels, overlapping stimuli were re-assigned to only a single level. This comparison is therefore likely to overestimate level differences for PD. Comparisons were performed using one-way ANOVA for scaled variables and the Kruskal Wallis Test for nominal/ordinal variables. Results are depicted in Table 6 - Table 9.

5.6.1 Matched characteristics

For all three naming tasks (PN, RN and LN), levels 1-4 were well matched for category (living/non-living and manmade/natural). Response length was well-matched for RN, LN and PD. Although response length differed statistically between levels for PN, this was not a meaningful difference; responses were only one letter shorter on average in level 1.

5.6.2 Between-level differences

For all tasks, level determinants differed significantly between levels: Stimuli had lower NA, a higher *H* statistic and a later AoA with increasing difficulty level. For PN, increasing difficulty levels also had lower word frequency. Between-level comparisons of pilot data (error rates and RT) confirmed both children and

adults found levels 1-4 progressively more difficult for all three naming tasks (PN, RN and LN), see Table 6 - Table 9. RTs did not differ between levels for PD. However, RT may not be a good measure of performance difficulty for a sentence-generation task (as discussed previously). Encouragingly, between-level differences were seen in two other measures of task difficulty for PD; response uncertainty (as measured by RA and *H*) and errors.

Picture Naming: Stimuli charactersitics for four difficulty levels									
	Level 1		Level 2		Level 3		Level 4		
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	*p*
Matched characteristics									
Response length (letters)	4.5	1.14	5.63	1.5	5.29	2.22	5.71	1.4	0.04
Living	25%	-	45.80%	-	54.20%	-	37.50%	-	0.21
Man-made	58.30%	-	33.30%	-	29.20%	-	41.70%	-	0.18
Level determinants									
Word frequency	6030	7404	6512	12155	2110	3197	1227	2033	0.03
Name agreement	1	0	0.97	0.05	0.88	0.12	0.8	0.11	<0.001
H	0	0	1.53	2.05	4	2.16	8.88	4.67	<0.001
AoA likert scale (age)	2 (<4 yrs)	-	3 (<6 yrs)	-	3 (<6 yrs)	-	3 (<6 yrs)	-	0.01
Pilot data (N=20)									
Child RT (ms)	1016.74	108.35	1010.63	99.12	1085.9	119.67	1184.63	177.18	<0.001
Adult RT (ms)	821.15	73.19	838.82	93.74	918.41	127.8	984.67	162	<0.001
Error rate (child)	<1%	1%	<1%	2%	<1%	1%	<1%	2%	0.68
Target group									
Age range	5-7 years		8-10 years		11-13 years		14+ years		N/A
Academic year	Reception - year 2		Year 3 - 5		Year 6 - 8		Year 9+		N/A

Table 6: Picture Naming (PN) stimuli characteristics.

Read and Name: Stimuli charactersitics for four difficulty levels									
	Level 1		Level 2		Level 3		Level 4		
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	*p*
Matched characteristics									
Response length (letters)	4	1	5	1	5	1	6	2	0.13
Stimulus length (words)	6	2	7	2	7	2	7	2	0.22
Stimulus length (characters)	25	9	28	10	29	9	31	9	0.13
Living / non-living	54% / 46%	-	46% / 54%	-	25% / 75%	-	25% / 75%	-	0.08
Natural / manmade	58% / 42%	-	50% / 50%	-	50% / 50%	-	62% / 38%	-	0.77
Level determinants									
Word frequency	4332	4566	5070	10335	4064	11894	3541	6570	0.41
Name agreement	0.95	0.05	0.88	0.09	0.80	0.10	0.64	0.22	<0.001
H	3.37	3.58	6.83	4.29	10.66	4.56	12.23	7.43	<0.001
AoA likert scale (age)	2 (<4 yrs)	-	2 (<4 yrs)	-	3 (< 6yrs)	-	3 (<6 yrs)	-	<0.001
Dictionary age									
Oxford First Dictionary (5-7 years)	92%	-	63%	-	33%	-	17%	-	
Oxford Primary Dictionary (7-11 years)	8%	-	25%	-	50%	-	54%	-	<0.001
Oxford Dictionary for Schools (11-16 years)	0%	-	13%	-	17%	-	29%	-	
Pilot data (N=20)									
Child RT (ms)	4034	762	4856	865	4953	999	5686	925	<0.001
Adult RT (ms)	1843	483	2294	698	2361	562	2627	747	0.001
Error rate (child)	3.13%	5.28%	5.83%	7.61%	11.67%	10.07%	16.04%	9.32%	<0.001
Target group									
Age range	5-7 years		8-10 years		11-13 years		14+ years		N/A
Academic year	Reception - year 2		Year 3 - 5		Year 6 - 8		Year 9+		N/A

Table 7: Read and Name (RN) stimuli characteristics.

Listen and Name: Stimuli charactersitics for four difficulty levels									
	Level 1		Level 2		Level 3		Level 4		
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	*p*
Matched characteristics									
Response length (letters)	5	1	5	1	5	1	5	2	0.47
Stimulus length (words)	7	2	7	2	7	2	8	2	0.30
Stimulus length (characters)	27	10	31	10	30	8	33	8	0.26
Living / non-living	54% / 46%	-	42% / 58%	-	25% / 75%	-	21% / 79%	-	0.20
Natural / manmade	62% / 38%	-	46% / 54%	-	54% / 46%	-	54% / 46%	-	0.57
Level determinants									
Word frequency	4424	7415	2981	4117	3453	5855	1204	1815	0.19
Name agreement	0.96	0.05	0.89	0.08	0.77	0.11	0.63	0.11	<0.001
H	3.05	3.97	6.56	4.30	11.37	4.79	15.58	4.34	<0.001
AoA likert scale (age)	2 (<4 yrs)	-	3 (< 6yrs)	-	3 (< 6yrs)	-	3 (< 6yrs)	-	0.002
Dictionary age									
Oxford First Dictionary (5-7 years)	75%	-	41%	-	29%	-	13%	-	
Oxford Primary Dictionary (7-11 years)	17%	-	42%	-	54%	-	73%	-	0.004
Oxford Dictionary for Schools (11-16 years)	8%	-	17%	-	17%	-	14%	-	
Pilot data (N=20)									
Child RT (ms)	4244	849	4898	898	5083	934	5674	990	<0.001
Adult RT (ms)	1868	431	2442	791	2453	650	2695	670	<0.001
Error rate (child)	3.13%	3.85%	9.17%	9.52%	12.92%	9.66%	13.96%	9.44%	<0.001
Target group									
Age range	5-7 years		8-10 years		11-13 years		14+ years		N/A
Academic year	Reception - year 2		Year 3 - 5		Year 6 - 8		Year 9+		N/A

Table 8: Listen and Name (LN) stimuli characteristics.

Picture Describing: Stimuli charactersitics for four difficulty levels									
	Level 1		Level 2		Level 3		Level 4		
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	*p*
Matched characteristics									
Min. response length (words)	6	1	6	1	6	1	7	1	0.17
Max. response length (words)	8	1	7	1	8	1	8	1	0.06
Living / non-living (both)	21% / 21% (58%)		8% / 21% (71%)		13% / 29% (58%)		13% / 33% (54%)		0.91
Natural / manmade (both)	29% / 8% (63%)		33% / 8% (58%)		33% / 21% (46%)		29% / 29% (42%)		0.79
Level determinants									
Response frequency	4377	4998	3703	6624	3623	6611	4067	5620	0.27
Response agreement	0.90	0.10	0.76	0.14	0.66	0.17	0.59	0.25	<0.001
H	4.44	3.65	9.38	4.59	12.66	4.84	14.92	10.35	<0.001
AoA likert scale (age)	3 (< 6yrs)	-	3 (< 6yrs)	-	3 (< 6yrs)	-	3 (< 6yrs)	-	0.66
Verb									
Eating	38%	-	25%	-	29%	-	21%	-	0.69
Drinking	8%	-	21%	-	33%	-	38%	-	
Jumping	37%	-	37%	-	9%	-	8%	-	
Falling	17%	-	17%	-	29%	-	33%	-	
Pilot data (N=20)									
Child RT (ms)	1480	190	1501	148	1487	177	1525	287	0.27
Adult RT (ms)	1282	168	1437	218	1432	215	1367	213	0.05
Child RD (ms)	6319	609	6512	819	6328	720	6656	1738	0.2
Error rate (child)	7%	11%	10%	9%	16%	9%	24%	14%	<0.001
Target group									
Age range	5-7 years		8-10 years		11-13 years		14+ years		N/A
Academic year	Reception - year 2		Year 3 - 5		Year 6 - 8		Year 9+		N/A

Table 9: Picture Describing (PD) stimuli characteristics.

6. DISCUSSION

Here I present norms for a new series of picture and sentence stimuli, which are more colourful, engaging, child-friendly and culturally valid relative to previous naming stimuli (Snodgrass & Vanderwart, 1980). These stimuli have been designed to aid accurate performance and encourage sustained task compliance in children as young as 5 years, and in children with epilepsy; who often suffer from comorbid attention difficulties (Hermann et al., 2007; Sherman et al., 2007). As such, I hypothesise the use of these stimuli in the new Panda Games task battery will enhance the success rate of language fMRI in healthy children and in children with epilepsy.

Only one other study could be found which has investigated naming responses and reaction times in children (Cycowicz et al., 1997). My findings emphasise the importance of using age-adjusted tasks when investigating language development; children performed less accurately and with more difficulty relative to adults. This finding has important consequences for interpreting differences in language network activation which have been found using a single task across a broad age range (Brauer & Friederici, 2007; Holland et al., 2007; Wilke et al., 2006). It also highlights the value of developmentally appropriate language tasks, which are surprisingly rare in neuroimaging investigations of language.

As a first step towards developing language tasks which accommodate for the wide range of ability levels in children and patient populations, I used RT data (a measure of effort) and error rates (a measure of accuracy) to establish how children can be grouped, in order to best explain between subject variance in task performance. I found that age is a better predictor of effort and performance accuracy than verbal ability; explaining 34% of variability in picture naming RTs, 50% of variability in RTs for naming items from a written or heard description (RN/LN), and 47% of variability in errors when describing a picture. Specifically, I found that grouping children into four age categories was optimal for explaining task performance; 5-7 years, 8-10 years, 11-13 years and 14-16 years. Assigning children to one of these four age groups explained 36%, 47% and 40% of variability in task performance for PN, RN/LN and PD,

respectively. Based on these findings, I aimed to develop four difficulty levels within each task, one for each age group.

I used established predictors of performance accuracy and performance difficulty from the naming literature (frequency, name agreement, *H* and AoA) to group stimuli into four levels; level 1 (very easy), level 2 (easy), level 3 (moderate) and level 4 (difficult). These levels were matched for stimulus category (living/non-living and natural/manmade), based on evidence from neuroimaging and lesion studies for different networks supporting category specific semantic processing (Farah, Mcmullen, & Meyer, 1991; Hillis & Caramazza, 1991; Leube, Erb, Grodd, Bartels, & Kircher, 2001; Okada et al., 2000). Between-level comparisons confirmed children found task performance more difficult (had slower RTs) and were less accurate (showed increased errors) on consecutive levels of each task. For PN, RT measures did not differ significantly between levels. This measure may have been less variable due to the automaticity of naming pictures, relative to naming from a description (RN/LN) or describing a scene (PD). Encouragingly, error rates did differ significantly between levels, suggesting that although children may find all levels of the PN task relatively easy (automatic), age-related differences are still apparent in the accuracy with which they are able to perform. Beyond appropriate task design, behavioural measures can also be used to reduce variability in performance difficulty and accuracy in children and patient populations, such as the provision of developmentally appropriate task instructions and extensive opportunity for practice. These methods are discussed in Chapter 3 (section 8.), where an optimised pre-scan protocol is outlined.

7. CONCLUSIONS

Performance on both receptive and expressive language tasks is more difficult and less accurate for younger children. In each of the Panda Games tasks, I developed four difficulty levels (very easy, easy, moderate and difficult) which children found increasingly more difficult to perform. These initial analyses suggest the Panda Games task battery has been well-designed to control for individual differences in task performance and effort. Using these tasks I aim to identify age-related changes in the language network, providing stringent control over task performance differences

(Chapter 8). I also aim to investigate whether the Panda Games provide more accurate mapping of core language regions, reducing the contribution of regions associated with sustaining effortful task performance (Chapters 5 and 6).

PART 2: METHODS

Chapter 3: Principal methods

1. AIMS

This chapter aims to summarise the main sample and methods used in the remaining chapters of this thesis. The sample of healthy children recruited for analyses reported in Chapters 4-8 (except Chapter 7) are described, along with an outline of the questionnaire measures and neuropsychological assessments administered during testing. Descriptives for children with epilepsy (Chapter 6) are reported in the relevant chapter, for ease of reference during reading. I also provide an overview of the main methods used throughout this thesis, including magnetic resonance imaging (MRI) and functional MRI. I introduce the biophysical principles underlying each method, the parameters used to obtain data presented in Chapters 4-8, the statistical preparation of fMRI data for within-subject and between-subject analyses, and how fMRI data were modelled. I also discuss how statistical inferences are made using fMRI data, highlighting important issues to be considered throughout this thesis. Finally, I outline a pre-scan preparation protocol and the Panda Games task battery, which was performed by all participants during scanning.

2. PARTICIPANTS

Forty three healthy children aged 5-16 years were recruited from local schools via poster advertisement (Table 10). Exclusion criteria were; 1) first language other than English, 2) visual impairment requiring correction with glasses, 3) auditory impairment requiring correction with a hearing aid, cochlear implant or other device, 4) history of developmental or neurological condition (including dyslexia, attention deficit hyperactivity disorder, autism spectrum disorder and preterm birth), and 5) metal implant in the body. Participants underwent neuropsychological assessment, fMRI scanning and were asked to complete a selection of questionnaire measures (described below). Two participants did not return for neuropsychological assessment and did not complete questionnaire measures.

3. NEUROPSYCHOLOGICAL ASSESSMENT AND QUESTIONNAIRE MEASURES

3.1 Socio-economic status

Both parents completed a socio-economic status questionnaire (Appendix 0) based on the five-class version of the self-coded National Statistics Socio-economic Classification System (NS-SEC: http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/soc2010/soc2010-volume-3-ns-sec--rebased-on-soc2010--user-manual/index.html). The NS-SEC is derived from occupation and employment status information. Parents also reported the number of years spent in education, as an additional measure of socio-economic status (Table 10).

3.2 Ethnicity

Participants or their parents reported their ethnicity (Table 10) using a tick box list based on the national standard for ethnic groups from the Department of Health (http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/Browsable/DH_5319155).

3.3 Handedness

Handedness was assessed using an age-appropriate *Family Handedness* questionnaire (Appendix 5.), and was assessed formally using the Edinburgh Handedness Inventory (Oldfield, 1971). The latter provides a handedness score, where scores of 40-100 indicate right handedness, scores of -40 to -100 indicate left handedness, and scores between -40 and 40 indicate ambidexterity (Table 10).

3.4 Intelligence

Intelligence was assessed using age-appropriate measures; the Wechsler Abbreviated Scale of Intelligence (Wechsler, 1999) for participants aged 6-16 years, and the Wechsler Preschool and Primary Scale of Intelligence third edition (Wechsler, 1989) for participants aged 5 years. The WPPSI-III shows strong correlation with measures of intelligence in older children (correlation with WISC-III subtests; r=0.69-0.89), although it has not been validated in comparison to the WASI specifically. Both

assessments provide measures of verbal IQ, performance (nonverbal) IQ and full scale IQ (Table 10). Standard scores have a mean of 100 and a standard deviation of 15.

3.5 Executive functions

Parents completed the Behavioural Rating Inventory of Executive Functioning (BRIEF), which is an 86-item questionnaire aimed at assessing executive functions in children aged 5-18 years (Gioia, Isquith, Guy, & Kenworthy, 2000). The BRIEF provides a Global Executive Composite (GEC) which summarises scores across eight clinical scales, including; inhibition, shifting, emotions, initiating, working memory, planning and organising, and monitoring. It also produces a Meta-cognition Index (MI) which reflects the child's ability to initiate, plan, organise, self-monitor, and sustain working memory, as well as a Behavioural Regulation Index (BRI) which measures the child's ability to shift cognitive set and also to modulate emotions and behaviour via appropriate inhibitory control. Standard scores have a mean of 50 and a standard deviation of 10 (Table 10). Higher scores reflect increased executive dysfunction.

3.6 Language

Participants were assessed using the Clinical Evaluation of Language Fundamentals – Fourth Edition UK (CELF-4^{uk}) (Semel, Wiig, & Secord, 2003b), suitable for children aged 5-16 years of age. This assessment provides a Core Language Score (a measure of overall language performance), a Receptive Language Index (a measure of listening and auditory comprehension), and an Expressive Language Index (a general measure of expressive language skills). All index scores have a mean of 100 and a standard deviation of 15, and are presented in Table 10.

	Min.	Max.	Mean	SD
Demographics				
Age	5	16	11	4
Edinburgh Handedness score	-100	100	52	58
Years in education (Mother)	15	38	23	5
Years in education (Father)	16	48	25	7
Neuropsychological performance				
WASI				
Full scale IQ	95	151	120	15
Verbal IQ	89	143	118	13
Performance IQ	91	149	117	17
CELF-IV				
Core language	96	133	111	11
Receptive Language	93	136	112	12
Expressive Language	93	140	114	12
BRIEF (Parent)				
General executive composite	36	68	49	8
Metacognitive index	36	73	49	9
Behavioural regulation index	37	66	48	7

	N	%
Sex		
Females	21	49%
Males	22	51%
Handedness		
Right	30	79%
Ambidextrous	4	11%
Left	4	11%
Socio-economic status		
Mother		
Manager and professional occupations	34	79%
Intermediate occupations	6	14%
Small employers and own account workers	0	0%
Low supervisroy and technical occupations	1	2%
Semi-routine and routine occupations	0	0%
Father		
Manager and professional occupations	32	74%
Intermediate occupations	0	0%
Small employers and own account workers	2	5%
Low supervisroy and technical occupations	5	12%
Semi-routine and routine occupations	0	0%
Ethnicity		
White (British, Irish, Other)	36	83.72%
Mixed (white and black, white and asian)	5	11.63%
Black or black British	2	4.65%
Number of languages spoken		
Monolingual	30	70%
Bilingual	10	23%
Trilingual	3	7%

Table 10: Descriptive statistics for a sample of 43 healthy children and adolescents aged 5-16 years. Scalar data are displayed in the table on the left, including the minimum (Min.), maximum (Max.), mean and standard deviation (SD) for the sample. Categorical data are displayed in the table on the right, including the number (N) and proportion (%) of participants. Neuropsychological assessments include the Wechsler Abbreviated Scale of Intelligence (WASI), Clinical Evaluation of Language Fundamentals (CELF-4^{uk}) and the Behaviour Rating Inventory of Executive Functioning (BRIEF-parent). Standard scores are reported for each assessment. For the WASI and CELF standard scores have a mean of 100 and standard deviation of 15. For the BRIEF standard scores have a mean of 50 and standard deviation of 10 (with higher scores indicating more dysfunction).

3.7 In-scanner experience

Participants were given a self-report *In-scanner experience* questionnaire at the end of their appointment, to complete on the way home or once they arrived home immediately after their appointment (Appendix 6.). This questionnaire was designed in collaboration with a Masters student (Lucy Palmer) to assess how participants felt during the brain scan. We designed the questionnaire to probe two themes shown to be of importance to scan success in the literature: physical comfort and researcher interaction (Thomason, 2009). Several questions were also included to assess 'task enjoyment', to see if experience with the Panda Games (e.g. boredom or difficulty) influenced scan results. Finally, we included a measure of in-scanner anxiety (a known cause of scan failure: Byars et al., 2002), based on the 6-item short form of the Spielberger State-Trait Anxiety Inventory (STAI) (Marteau & Bekker, 1992). All questions were designed on a 4-point likert scale using developmentally appropriate language, with smiley/sad faces beneath each point of the likert-scale for children aged 5-7 years, to help them complete the questionnaire accurately (see Appendix 6. for examples). All items were scored with a value between 1-4, with positive responses scoring higher values. Twenty four children (56% of the sample) returned completed questionnaires. Factor extraction was performed using Principle Component Analysis to identify the underlying factor structure of this questionnaire (Appendix 7.). Three factors were extracted, reflecting; 1) *In-scanner experience* (including physical comfort and task enjoyment), 2) *Researcher interaction*, and 3) *Headphone comfort*. Composite scores for each factor were based on the mean score for each individual. These scores were used as predictors of data quality (in-scanner movement) in Chapter 5.

4. MAGNETIC RESONANCE IMAGING (MRI)

4.1 Principles

How the MRI scanner works and the biophysical underpinnings of the MR signal (Gadian, 1995; Hornak, 1996) are described below, and are also shown in Figure 10 and Figure 11.

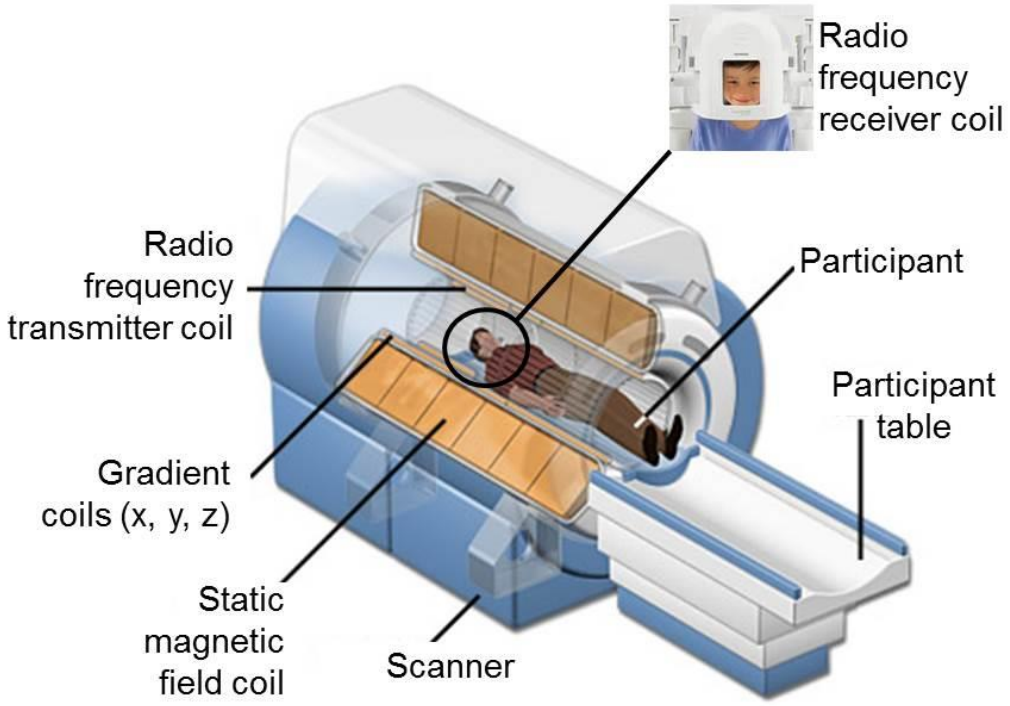


Figure 10: Cross-section of a Magnetic Resonance Imaging (MRI) scanner. The main components of the MRI scanner are shown in relation to a human participant.

4.1.1 Water, protons and axes of spin

Around 80% of the human body is made up of water. The Hydrogen atom is abundant in water, and has magnetic properties. It is therefore the ideal candidate for obtaining a magnetic resonance (MR) signal in humans. Crucial to understanding how an MR signal is obtained from hydrogen nuclei, is the concept of spin. Spin is a property of hydrogen atoms (hereon referred to as protons), which can be visualised as the spin of a proton around its own axis (Figure 11 A). Spin creates a magnetic field (*M*) for each individual proton. Outside of the MRI scanner, axes of spin (and therefore *M*) occur in random directions (Figure 11 A). However, because protons are magnetic, *M* can be adjusted through the use of magnets.

4.1.2 The static magnetic field (B_0)

The MRI scanner is composed of several magnets. The largest (the static magnetic field coil) forms the main part of the scanner, within which the participant lies (Figure 10). This magnet creates a static magnetic field (B_0) which is very strong (1.5 tesla); even stronger than the magnet in a car scrap yard (1 tesla). This magnetic field is constantly on, unlike some of the other magnets in the scanner. When a participant lies within the scanner bore, *M* aligns with B_0 (Figure 11 A). No MR signal can be detected when protons are in this low-energy state.

4.1.3 A transient magnetic field (B_1)

The radio frequency (RF) transmitter coil (Figure 10) produces a second magnetic field (B_1) which can be switched on and off. This magnetic field is at right angles to the static magnetic field (B_0). When it is turned on, B_1 acts as an external force which causes precession of *M* between the two magnetic fields (like the wobble of a spinning-top). During precession protons are in a high energy state, which eventually causes *M* to flip. With a short burst of excitation from the RF transmitter coil *M* flips 90^0, to align with B_1. Protons then spin within the *xy* plane (Figure 11 B). When the excitation is prolonged *M* flips 180^0; causing an inversion relative to its low energy state (i.e. when aligned with B_0). Once the axis of spin has inversed the RF transmitter is switched off. The protons begin to return to a low energy state, and re-align with B_0.

4.1.4 Localising the MR signal in 3D space using Gradient coils

Once the RF transmitter coils are switched off, three gradient coils are switched on (Figure 10). These coils are crucial for MRI acquisition in humans, as they enable detection of the MR signal in 3D space. They do this by introducing gradients in the strength of the magnetic field, which alters the rate of precession in three different directions; x, y and z. The x, y and z gradient coils are switched on in various sequences and combinations, and together help locate the origin of the MR signal in 3D space.

4.1.5 Receiving the MR signal

The radio frequency receiver coil is placed in proximity to the part of the body being investigated. For neuroimaging, the RF receiver coil is placed around the head (i.e. the head coil) (Figure 10). The head coil records the MR signal emitted by protons as *M* realigns with B_0 and protons return to their low energy state (a process known as relaxation). The time it takes for a proton to relax to its low energy state, and the extent of relaxation which occurs before the next RF transmission pulse begins, produces variation in the MR signal. Crucially, because protons in different tissue classes have different relaxation rates, recording the MR signal at different time points allows the identification of MR signal from specific tissue classes.

4.1.6 Relaxation times and tissue classes

When protons relax back to their low energy state (when *M* realigns with B_0) they emit the MR signal, which is recorded by the RF receiver coil. MR signal can be recorded from two types of relaxation process (Figure 11 C), which occur at different time points following cessation of B_1. These also occur at different rates for grey and white matter, allowing identification of different tissue classes. Spin-lattice (or longitudinal) relaxation reflects the time it takes for signal to increase in the direction of B_0 (the z plane): Specifically, the time required for 63% of *M* to realign with B_0. Spin-spin (or transverse) relaxation reflects the time it takes for signal to decay in the xy plane (in the direction of B_1). Specifically, the time required for signal in the xy plane to decay back to 37% of its previous strength. Spin-lattice relaxation occurs early after B_1 ends (known as time 1). Spin-spin relaxation occurs more slowly following cessation of B_1 (at time 2). White matter produces the highest signal on T1 images, while CSF produces the highest signal on T2 images (Figure 11 D). These images are crucial for identifying structural abnormalities in clinical populations (particularly T2). The T1 image is crucial for neuroimaging investigations of language presented in this thesis, to allow localisation of function to underlying brain anatomy.

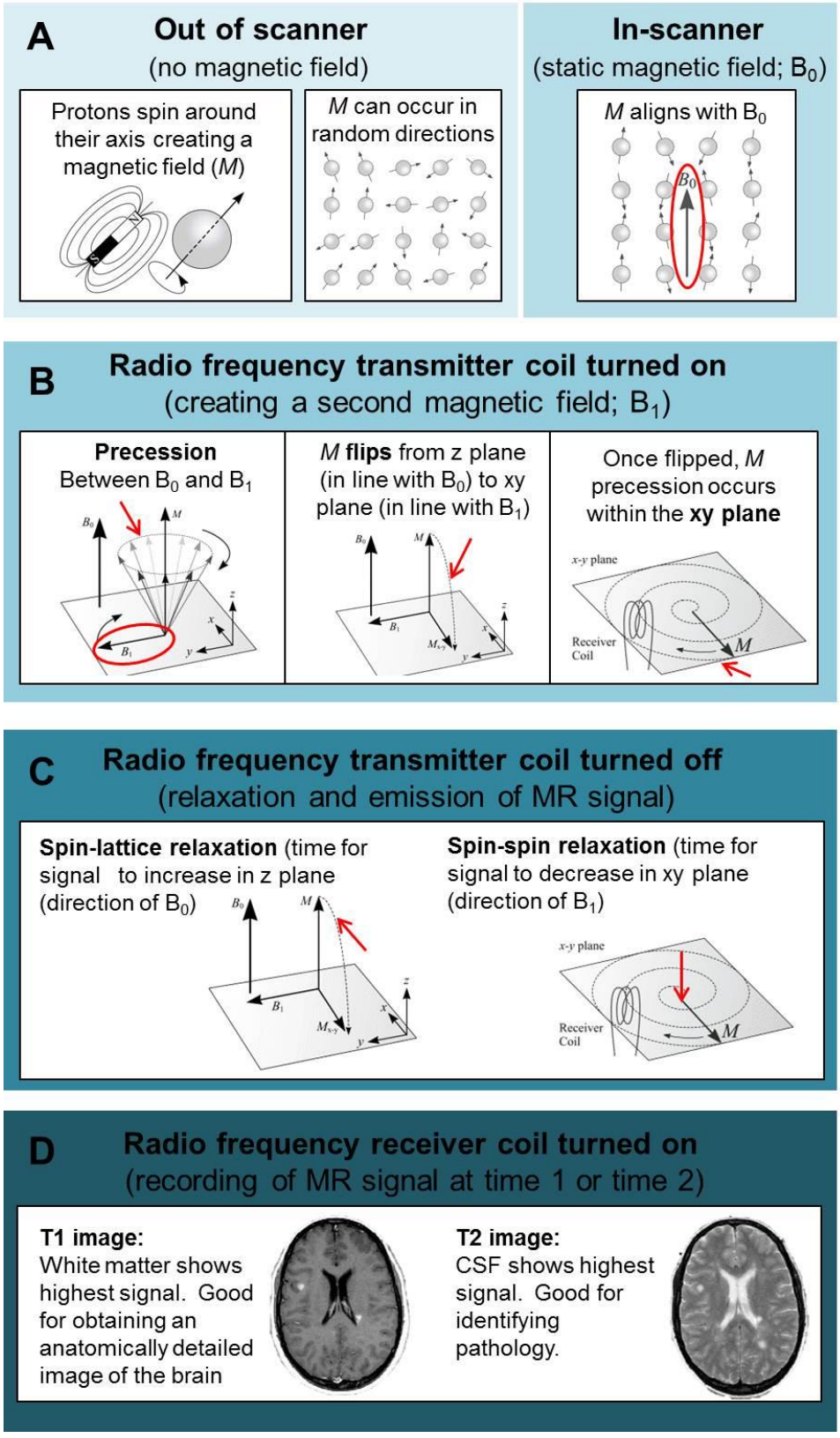


Figure 11: From proton to picture: The biophysics of the magnetic resonance (MR) signal. The behaviour of hydrogen atoms (protons) are shown for each stage of MRI scanning, from the participant entering the scanner (A), to excitation of protons during

scanning (B), recording of the MR signal (C) and the resulting images (D). To create this schematic, I used proton images taken from http://www.wikidoc.org/index.php/Basic_MRI_Physics.

4.2 Acquisition parameters

All participants were scanned with a 1.5 Tesla Siemens Avanto System (Erlangen, Germany). Three-dimensional data sets were acquired using a T1-weighted 3-dimensional fast low angle shot (3D-FLASH) sequence (TR = 11 milliseconds (ms), TE = 4.94 ms, flip angle = 15°, field of view = 256mm, matrix size = 256 x 256). This anatomical scan served to screen for anatomical abnormalities and was used to localise functional MRI results on an individual basis.

5. FUNCTIONAL MRI (FMRI)

5.1 Principles

Functional MRI works on similar principles as MRI, except it measures variation in the MR signal caused by the haemodynamic response of brain tissue to stimuli, as opposed to the relaxation of hydrogen atoms. The haemodynamic response and the resulting Blood Oxygen Level Dependent (BOLD) signal measured by fMRI, are discussed below.

5.1.1 Neurovascular coupling

Neural events in the brain are paralleled by a haemodynamic response which aims to meet the increased demand for glucose and oxygen from active neurons (Leniger-Follert & Hossmann, 1979). The co-occurrence of this haemodynamic response with neuronal activity is known as neurovascular coupling. The haemodynamic response reflects – in aggregate – changes in three aspects of cerebral dynamics; cerebral blood flow, blood volume and blood oxygenation (Ogawa et al., 1993). Changes in either of these can therefore affect fMRI findings. However, statistical modelling used in the analysis of fMRI results assumes a similar haemodynamic response across individuals, regardless of cerebrovascular differences. The onset of the 'typical' haemodynamic response is delayed by ~2 seconds following stimulus onset (Kwong et al., 1992). The

increase in blood flow actually overcompensates for the increase in oxygen demand, resulting in an oversupply of oxygenated blood at its peak (Fox & Raichle, 1986, 1988). The HRF reaches a plateau after 6-12 seconds, and then returns to baseline over a further 6-12 seconds. A post-stimulus undershoot is often evident (Frahm, Kruger, Merboldt, & Kleinschmidt, 1996), see Figure 12.

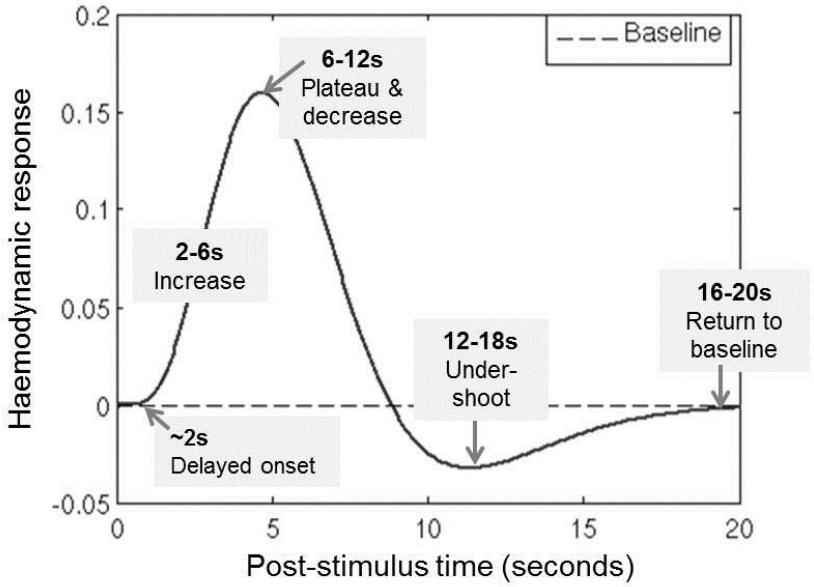


Figure 12: The haemodynamic response to a stimulus. Stimulus onset (0 seconds) is followed by an increase in oxygenated blood from approximately 2-6 seconds, followed by plateau (6-12 seconds), and an eventual decrease back to baseline (over 6-12 seconds), often followed by an undershoot (~12-15 seconds). The haemodynamic response takes ~16 seconds in total to return to baseline.

5.1.2 The Blood Oxygen Level Dependent (BOLD) signal

The Blood Oxygen Level Dependent (BOLD) signal was first described by Seiji Ogawa in the study of rat brains (Ogawa & Lee, 1990; Ogawa, Lee, Kay, & Tank, 1990; Ogawa, Lee, Nayak, & Glynn, 1990), and in 1992 the first studies of human cognition using fMRI (without contrast agents) were published (Bandettini, Wong, Hinks, Tikofsky, & Hyde, 1992; Kwong et al., 1992; Ogawa et al., 1992). The BOLD signal reflects local magnetic field inhomogeneities caused by the

haemodynamic response to neural activation. The over-compensatory haemodynamic response increases the ratio of oxyhaemoglobin to deoxyhaemoglobin in the vasculature surrounding active neurons (Malonek & Grinvald, 1996). Because deoxyhaemoglobin is more strongly paramagnetic than oxyhaemoglobin, changes in the ratio of oxygenated and deoxygenated blood changes the local magnetic field strength. Local magnetic field inhomogeneities introduce variance in spin-spin relaxation times (Figure 11 C), which can be recorded at a specific time point (T2*) by the RF receiver coil. The variation in spin-spin relaxation caused by these local field inhomogeneities are the source of BOLD signal. Essentially, we infer neuronal activity from this indirect measure of oxygen supply in the brain.

5.1.3 The neurophysiological basis of BOLD: Input, intra-cortical processing, or output?

Neuronal activity can be measured directly in animals using extracellular electrodes. There are three types of direct recording I will discuss here, each capturing a specific type of neuronal process; input, local (intra-cortical) processing or output (Logothetis, 2003). Extracellular field potentials (EFP) are measured by single microelectrodes, and reflect spiking activity (action potentials) of a single cell. EFP's therefore represent cell output; with larger EFPs correlating with larger cell bodies. Multiple Unit spiking Activity (MUA) can be measured when the electrode is moved further from the source of the spiking activity, and reflects the weighted sum of spiking activity; output from a group of neurons. The MUA can be low pass filtered (at <300Hz) to extract a low frequency local field potential (LFP). The LFP is a measure of slow events occurring within a group of neurons, including; synaptic potentials, afterpotentials of somatodendritic spikes, and voltage gated membrane oscillations. These processes reflect input to a cortical area, as well as local intra-cortical processing (including activity of excitatory and inhibitory interneurons). The relationship of these direct neurophysiological measures (EFP, MUA and LFP) to the BOLD signal, can help elucidate which neuronal processes are being measured by fMRI. In monkey cortex, EFPs are a poor predictor of the BOLD response, while MUA and LFP are significantly better predictors (Logothetis, Pauls, Augath, Trinath, &

Oeltermann, 2001). The best predictors are LFPs however, explaining ~8% more variance in the BOLD response than MUA (Logothetis et al., 2001). Further, while spiking (as measured by EFPs and MUA) returns to baseline just 2.5 seconds after stimulus onset, the slow wave LFPs continue for the duration of a stimulus, similar to the BOLD response (Logothetis, 2003). These findings suggest the BOLD response directly reflects neuronal processing elicited by a stimulus; principally, slow wave dendritic and integrative activity (system input and local intra-cortical processing), rather than axonal firing (system output). These processes are primarily associated with pyramidal cells in layer IV of the cortex (Logothetis, 2003). Whether the BOLD response exclusively reflects excitatory or inhibitory processes (or both) however is not yet known (Logothetis, 2008).

5.1.4 Developmental changes in the BOLD response

Many cerebro-vascular changes occur throughout childhood and adolescence which may impact the BOLD signal, including changes in neurovascular coupling and neuronal energy consumption (Harris, Reynell, & Attwell, 2011). If not taken into account, it is possible maturation of the haemodynamic response may confound findings of age-related changes in network activation, thought to reflect changes in neuronal information processing. Encouragingly however, initial studies suggest major differences in the peak latency of the BOLD response only occur early in life (from 0-2 years), with no significant change from 6-16 years of age (Jacobs et al., 2008; Kang, Burgund, Lugar, Petersen, & Schlaggar, 2003). Further, I will provide statistical control over age-related changes in the onset and duration of the BOLD signal (see section 9.2).

5.2 Acquisition parameters

Functional data were acquired using a 1.5 Tesla Siemens Avanto System (Erlangen, Germany), equipped with a 12 channel head coil. Blood oxygen level-dependent (BOLD) changes were measured using a whole brain echo-planar imaging (EPI) sequence (TR = 2160 ms, TE = 30 ms, flip angle = 75°, FoV = 210 mm, slice thickness = 3mm, 1mm inter-slice gap, slices = 30, voxel size = 3.3 x 3.3 x 3 mm^3).

The first two volumes were discarded as dummy scans to allow for magnetic saturation effects, resulting in 106 volumes per task. Axial images were collected parallel to the anterior commissure–posterior commissure plane, which served as an origin of reference. Stimuli were presented using a Windows PC and Cogent 2000 software; developed by the Cogent 2000 team at the Wellcome Trust Centre for Neuroimaging (WTCN) and the UCL Institute of Cognitive Neuroscience (www.vislab.ucl.ac.uk). Stimuli for *The Panda Games* were presented using scripts written by Dr Tom Hope and Dr Oiwi' Parker at the WTCN, in collaboration with Professor Cathy Price. Auditory stimuli were transmitted through MR compatible headphones equipped with an active noise cancellation system and speech responses were transmitted using a sensitive head-mounted microphone (http://www.mr-confon.de/en/technology.html). Visual stimuli were presented via an MR-compatible wall-mounted screen behind the scanner, and were viewed through a mirror fixed to the head coil.

6. STATISTICAL PRE-PROCESSING OF FMRI DATA

In order to obtain results from fMRI, data from the scanner undergo several stages of statistical pre-processing which aim to identify changes in the BOLD signal taking into account the shape of the haemodynamic response (Figure 12), and while controlling for sources of artefact. These stages of data preparation act to increase the signal to noise ratio (SNR), resulting in more power to detect effects of interest. They also prepare the data for within-subject and between-subject comparisons. Each stage is discussed below, and is based on information provided in the spm8 manual (http://www.fil.ion.ucl.ac.uk/spm/doc/spm8_manual.pdf), online course materials provided by the UCL Functional Imaging Laboratory (http://www.fil.ion.ucl.ac.uk/mfd/) and specific methods papers (referenced throughout).

6.1 Slice timing correction

When slices are acquired in sequential order, slices within a volume are acquired at different times. For example, for an ascending sequence (used in this thesis), slices at the bottom of the brain are acquired earlier than slices toward the top of the brain,

resulting in a time-shifted BOLD response in some regions of the brain. Slice timing correction accounts for differences in the acquisition time of slices within a volume, and increases sensitivity to detect activation (Sladky et al., 2011).

6.2 Realignment (within-subject realignment)

Realignment corrects for head movements which occur between volumes. During this stage each volume is aligned with a reference scan (the mean), using a rigid body transformation (Friston, Frith, Frackowiak, & Turner, 1995). Rigid body transformation only allows translations (movements in the x, y or z direction) and rotations (pitch, roll and yaw) of the volume (Figure 13). A realignment parameter file is produced, describing these translations and rotations. Essentially, this file contains information about how much the participant moved over time during

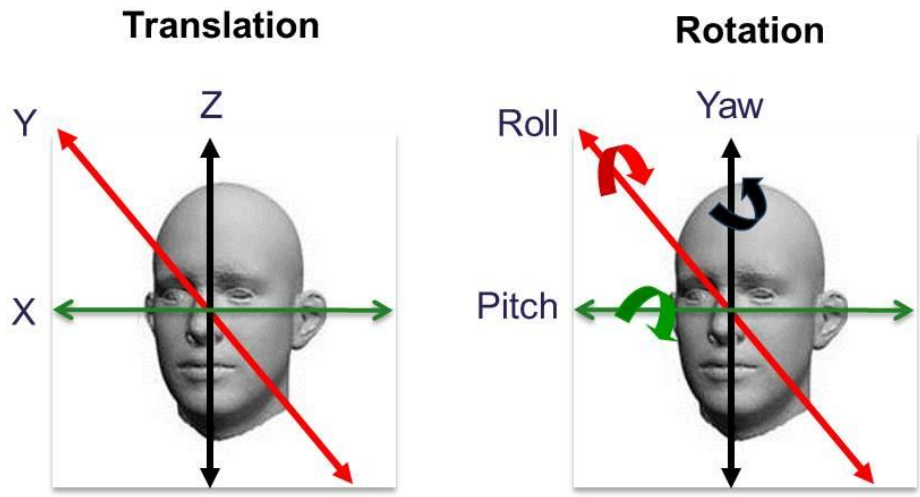


Figure 13: Rigid body transformation during image realignment. All functional scans are aligned with a template (the mean) using transformations in three directions (*x*, *y* and *z*) and rotations in three directions (pitch, roll and yaw).

scanning, and is used later in the pre-processing pipeline, to provide statistical control over in-scanner movement. Once realignment has been estimated, volumes are re-sliced using a B-Spline interpolation method, such that all voxels are aligned with the reference image (Unser, Aldroubi, & Eden, 1993a, 1993b).

6.3 Segmentation of the T1 scan

Unified segmentation uses a generative model to segment the T1 scan into three tissue classes, and uses the information from each tissue class to optimise warping parameters used later, to normalise both structural and functional scans to a standard template (Ashburner & Friston, 2005). The T1 is segmented into three different tissue classes; grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF) (Figure 14). Voxels are classified as GM, WM or CSF depending on the voxel value (intensity of the MR signal) and depending on the voxel's location within the brain volume. The likelihood of a voxel at any given location belonging to a specific tissue class is provided by Tissue Probability Maps (TPMs) (http://www.loni.ucla.edu/ICBM/ICBM_Probabilistic.html). These TPMs are in Montreal Neurological Institute space (MNI, ICBM152); a standardised space based on the average of 152 normal MRI scans from the International Consortium for Brain Mapping (Figure 14). For TPMs to guide the segmentation, the T1 is first warped (normalised) to the same standard space as the TPMs; this process is also optimised using information from each of the tissue classes (based on small-scale neuroanatomy). Segmentation therefore produces a parameter file which can be used to warp the T1 scan – and any other image co-registered to the T1 scan - from native (individual) space to MNI space. The unified segmentation procedure also corrects for biases in the magnetic field during scanning, which produce spatially smooth signal inhomogeneities across the volume. The bias-corrected T1 contains more uniform voxel intensities within each tissue class compared to the original T1. The bias corrected T1 was therefore used in all further analyses. Segmentation can be problematic in patient populations, due to the presence of structural lesions whose voxel values do not conform to the TPMs. However this was not a problem for my patient sample (Chapter 6); all tissue classes were segmented accurately.

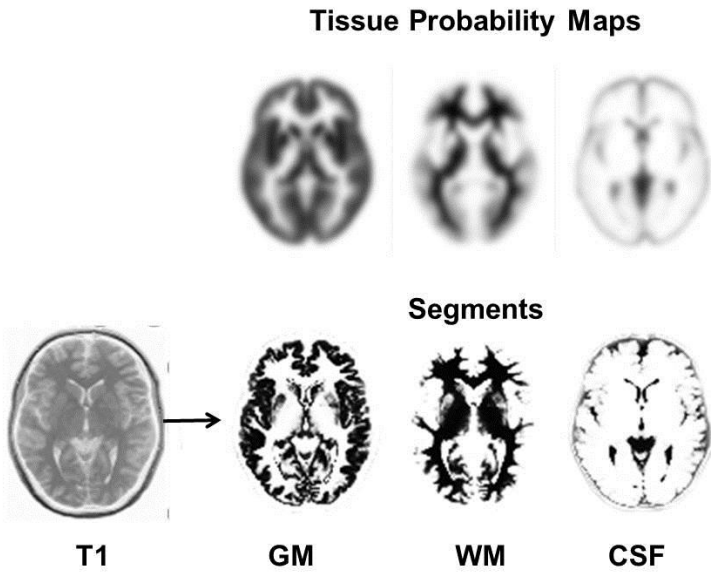


Figure 14: Segmentation of the T1 using Tissue Probability Maps.

6.4 Co-registration (inter-modal registration)

Co-registration aligns different image types to the same space, within an individual (Collignon et al., 1995). Functional images (the source images) are aligned with the T1 scan (the reference image) to allow anatomical localisation of single subject activations. In spm8, co-registration works by maximising the mutual information between the two different types of image using a 12 parameter affine transformation. This transformation not only allows translations and rotations (similar to realignment), but also zooms and shears (Figure 15 A). Functional scans in this thesis were co-registered and resliced to match the bias-corrected T1. This stage produces two histograms, showing the mutual information between modalities before and after registration (Figure 15 B). Voxel intensities in the source and reference images are shown on the *x* and *y* axes. After co-registration the histogram should be sharper; indicating greater agreement in voxel values (increased mutual information) between the two image types.

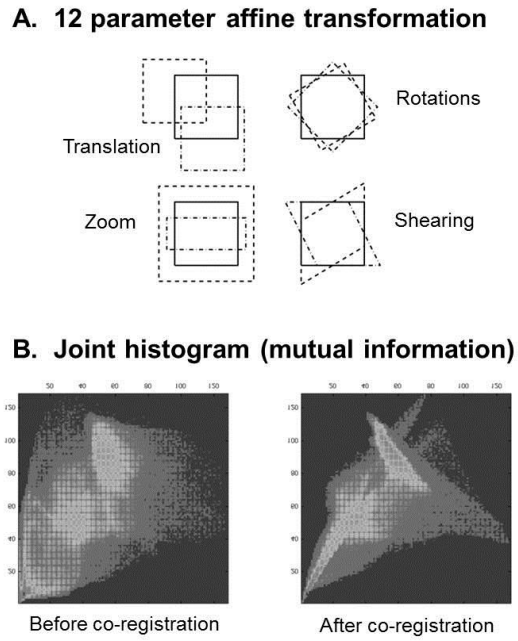


Figure 15: Co-registration of structural and functional scans using a 12 parameter affine transformation (A) to increase the mutual information between images (B).

6.5 Normalisation (between subject realignment)

Because individual brains vary in shape and size, fMRI data must be warped into standard space, such that results can be compared between subjects. Normalisation is performed using the warping parameters produced during the segmentation stage. Some studies normalise paediatric data to a custom-made, group-specific template (Ashburner, 2007). However, utilising adult-based atlases for registration of paediatric data has been demonstrated to be valid for children as young as at least 7 years of age (Burgund et al., 2002; Kang et al., 2003). This also provides the benefit of interpreting findings in relation to previous findings also reported in MNI, and allows the use of automatic anatomic labelling (Tzourio-Mazoyer et al., 2002).

6.6 Smoothing

Smoothing is performed to improve the signal-to-noise ratio (by reducing the effect of random voxel-level noise in the data) and to control for between-subject differences in neuroanatomy not fully accounted for by the normalisation stage. During smoothing, the intensity of each single voxel is blurred using a Gaussian

kernel, which creates a weighted mean intensity based on neighbouring voxels within a 3D spherical region of interest. Outer voxels in the spherical region of interest contribute least and immediate neighbours contribute most to calculation of the mean. The full-width at half-maximum (specified in the *x*, *y* and *z* directions) determines the size of this gaussian kernel. It is recommended that a FWHM of between 8,8,8 and 10,10,10 is used for group-level analyses: I used a FWHM of 8,8,8.

6.7 First level analyses (modelling within-subject variance)

The intensity of the BOLD signal in a given voxel over the course of the scanning session is known as the time series. During first-level analyses, the General Linear Model is used to predict the timeseries of each voxel (*y*) using a linear combination of predictor variables (x_1, x_2, $x_{3...}$) and an error term (*e*) (Figure 16) (Friston, Holmes, et al., 1995). The General Linear Model is therefore performed thousands of times, on a matrix of statistics in 3D space. Correction for multiple comparisons is therefore very important, and is discussed in section 7.1 .

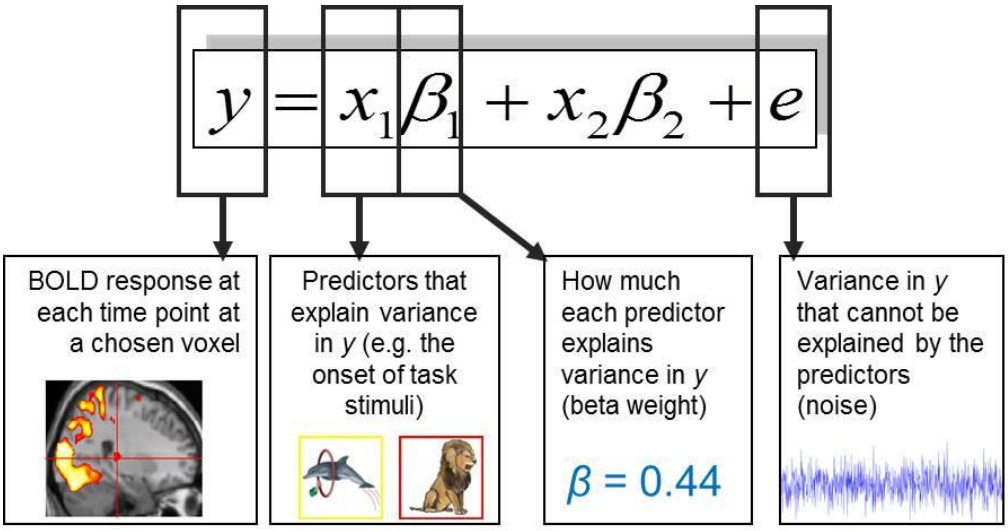


Figure 16: The General Linear Model. Equation used to estimate the optimal beta weights for each predictor variable in the design matrix for every voxel in the brain, to explain variance in *y* and reduce *e*.

6.7.1 Design matrix

All known predictor variables are entered into a design matrix (Figure 21). Predictor variables include effects of interest on the BOLD response (i.e. performing a cognitive task) and effects of no interest (e.g. the overall mean signal and in-scanner movement). The effect of performing a cognitive task is modelled using a stimulus function (or combination of stimulus functions) which specifies the onset and duration of individual stimuli (event-related design), or the onsets and durations of blocks of multiple stimuli (blocked design). Each stimulus function is convolved with a canonical model of the haemodynamic response (Figure 17 A) to account for the temporal profile of the BOLD response, and to model observed changes in y more accurately (K. J. Friston, Frith, Turner, & Frackowiak, 1995; K. J. Friston, A. P. Holmes, et al., 1995). Differences in the onset or duration of the BOLD peak can also be modelled (and therefore controlled) by including the temporal and dispersion derivatives of the canonical haemodynamic response, respectively (Figure 17 B) (Henson, Rugg, & Friston, 2001; Josephs, Turner, & Friston, 1997). In-scanner movement is typically modelled in the first-level design matrix as an effect of no interest, using the parameters produced during the realignment stage. These aim to control for movement effects which remain despite realignment, and to control for interpolation errors introduced by the realignment process (Friston et al., 1996). A constant term is also included in the design matrix, to control for mean (overall) signal strength. Before the model is estimated, data are high pass filtered to remove slow drifts in signal strength which occur over time during scanning (Figure 17 C).

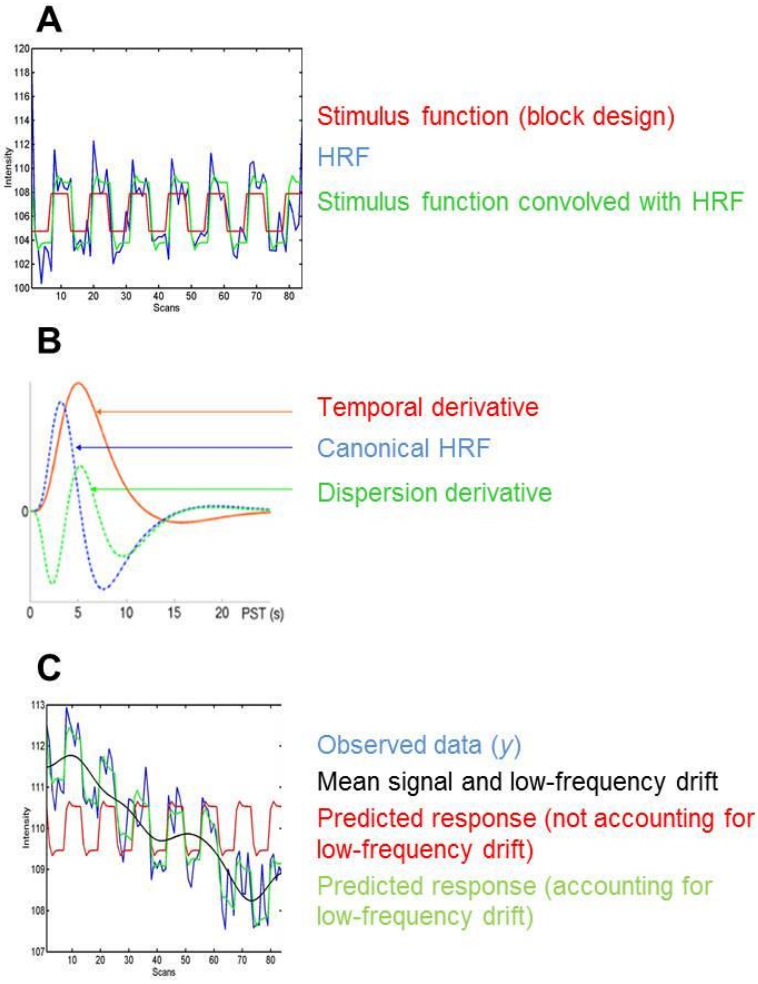


Figure 17: Modelling the Haemodynamic Response Function (HRF). The stimulus function is convolved with a canonical model of the HRF (A), temporal and dispersion derivatives are also model to account for variation in the onset and duration of the HRF (B), and the stimulus function is adapted to take into account mean signal (the constant) and low frequency signal drifts which occur throughout the scanning period (C).

6.7.2 Model estimation

Once the design matrix has been specified with all known predictors variables, the General Linear Model is estimated using ordinary least squares, to provide beta weights (β) for each predictor variable. The beta weights are calculated to quantify how much each predictor contributes to observed changes in the time series (y), with the aim of reducing e (unexplained variance in y).

6.7.3 F and t contrasts

The null hypothesis is tested using either *F* or *t* tests. The former tests the null hypothesis that beta weights are 0 for any of the specified predictor variables, and produces an *F* statistic. The latter tests the null hypothesis that a specific combination of beta weights are 0, and produces a *t* statistic. The desired combination of beta weights is specified using contrast vectors, in which each row represents a contrast and each vector specifies regressors in the design matrix (conditions). For example, a *t* contrast of [1 -1] would test the null hypothesis that there is no increase in beta weights for condition 1 versus condition 2. The beta weights associated with each contrast are saved as contrast images, which are used in the second-level (group) analyses.

6.8 Second-level analyses (modelling between-subject variance)

Contrast images from the first-level analysis for each individual participant are entered into a second-level random effects analysis. Because random effects analysis takes into account the between-subject variance (as opposed to fixed effects analyses, which treat subjects as fixed factors), results can be generalised from the sample to the population. Contrast vectors are used to specify contrasts of interest, as at the first level, and produce statistical parametric *F* or *t* maps.

7. STATISTICAL INFERENCE

Once fMRI results have been estimated at the first or second level, they can be viewed over several statistical thresholds. The issue of statistical thresholding is an important one for fMRI data, and still receives much discussion and debate. The main points of this debate and potential solutions are outlined below, as well as the strategy I have adopted in this thesis to address the issue of statistical thresholding. Results can also be reported at different levels of inference; whether peaks, clusters or whole sets of clusters are significant. The conditions under which each level of inference is most valid are discussed in the sections below, and the most suitable level of inference for this thesis is highlighted.

7.1 Statistical thresholding and control of type I and II errors

When spm *t* map results are viewed, inferences are made based on the topology of the test statistic *t* in space (its height and its extent) and over time. Reporting of regional BOLD signal change depends on the statistical threshold applied (α), above which the test statistic (*t*) is considered statistically significant (Figure 18).

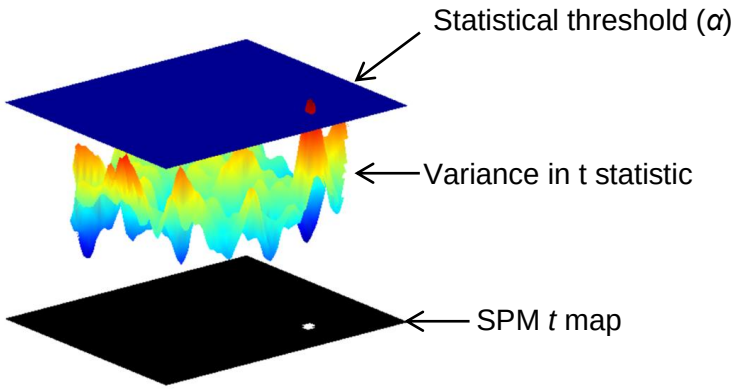


Figure 18: Example of applying a height threshold (blue square) to topographical data, and the resulting region of significant activation on an spm *t* map (white region on the black square).

The statistical threshold has an associated *p* value, which reflects the chance of *t* if the null hypothesis (H_0) is true (the null being that observations are the result of random chance). A *p* value of 0.05 therefore suggests 5% of observations could be due to random chance alone. These 5% are known as false positives, and reflect a type I error (false rejection of the H_0). Although a *p* value of 0.05 is commonly used for univariate statistics (e.g. behavioural studies), fMRI is based on multivariate statistics; with statistical tests being performed on a voxelwise basis in the order of tens of thousands. For example, if there are 50,000 voxels in a whole-brain spm *t* map, this reflects 50,000 *t*-tests. In this case, a *p* value of 0.05 would indicate that 2500 of the observations (i.e. activation peaks) could be type I errors. This can have fatal consequences for the interpretation of fMRI results (Bennett, Miller, & Wolford, 2009). For neuroimaging studies, the statistical threshold therefore needs

to be adjusted (corrected) to account for the number of comparisons made; thus reducing the type I error rate. Corrections can be made using the Bonferroni procedure (α /number of comparisons performed) to control for the Family Wise Error (FWE) rate; where H_0 = the *whole family* (volume) of voxels arose by chance. It is important to note that Gaussian Random Field Theory - implemented in spm8 – adjusts this calculation to account for spatial smoothing of data, by calculating the rate of false positive peaks rather than voxels. This method produces results with high specificity (few type I errors) but reduces sensitivity, resulting in increased type II errors (failure to reject the null).

Lieberman and colleagues argue that focusing on controlling type I errors and ignoring the concurrent inflation of type II errors introduces a bias towards large and obvious effects (i.e. in sensory and motor systems) rather than more subtle effects often seen for higher order cognitive functions or social and affective neuroscience, where intra- and inter- subject variability can be high (Lieberman & Cunningham, 2009). The authors further argue that type II error rates should be of more concern than type I error rates because type I errors (which occur randomly throughout the spm *t* map) are self-correcting on replication and through data aggregation (i.e. meta-analyses). For example, should Bennett have scanned more than one salmon, it is unlikely false positives would have occurred in the same region across the whole shoal. In contrast, false negatives cannot be replicated or identified in meta-analyses, because insignificant or trend-level findings are typically unreported. As such, the inflation of type II errors associated with correcting for FWE may cause real effects to be overlooked. When deciding on which statistical threshold to use, it is also important to consider the potential effects of making a type I or a type II error in the context within which inferences are being made. For example, when fMRI is used in an exploratory and experimental manner to generate hypotheses, making a type II error might be of less concern than making a type I error (although Lieberman and colleagues would disagree). In contrast, when fMRI is being used to inform neurosurgery, making a type II error can have serious consequences. In this instance accepting a higher type I error rate may be more prudent (see Chapter 6, Patient A). For these reasons I have investigated alternative thresholding methods which aim to provide a more balanced control over both type I and type II error rates.

One option, used in early neuroimaging studies, is to adopt a combined intensity (e.g. $p<0.001$) and minimum cluster size threshold (e.g. $k>10$) (Forman et al., 1995). Alternatively, one can adjust α based on the False Discovery Rate (FDR) (Benjamini & Hochberg, 1995). The FDR reflects the *proportion* of all positive findings which could be false, and provides increased sensitivity relative to FWE with minimal cost in terms of false positives (Chumbley, Worsley, Flandin, & Friston, 2010). As such, using an FDR of 0.05 may represent a more ideal balance of specificity and sensitivity in fMRI investigations of higher-order cognition, relative to FWE. Unlike earlier versions, spm8 controls for the FDR of peaks or clusters (Chumbley & Friston, 2009), rather than voxels (Genovese, Lazar, & Nichols, 2002). This has been shown to provide better control of type I errors (Chumbley et al., 2010). However, visualisation of results with FDR correction is not yet implemented in spm8, limiting its application.

Having reviewed methods for statistically thresholding neuroimaging data, for all chapters hereafter I report statistics corrected for FWE (when $p<0.05$) within a whole-brain grey matter template provided in spm8 (to exclude true negatives - e.g. in CSF and white matter – and reduce multiple comparisons; as recommended by Genovese et al., 2002). Where I had apriori hypotheses, a small volume correction was applied (Worsley et al., 1996), and is discussed in the relevant chapter. Further, in response to the claim that failing to report neuroimaging trends prohibits the advantage of meta-analyses to identify subtle effects (otherwise obscured by small sample sizes) I also report results at a more lenient combined intensity and cluster size threshold ($p<0.001$, $k>10$). For the patient pilot study more lenient thresholds were used and are discussed in the Chapter 6.

7.2 Topological inference

The inferences that can be made depend on which topological feature of the statistical parametric map is reported. Peak-level inferences can be made when the *height* of an individual peak is higher than would be expected by chance. Cluster-level inferences can be made when regions with peaks over a height threshold are

more extensive than expected by chance. Set-level inferences can be made when the set of clusters (determined by *extent and height* thresholds) is larger than the set of clusters that would be expected by chance. Each increasing level of inference has increasing statistical power (sensitivity), but with decreasing control over false positives and therefore decreasing localisation (Friston, 2007), see Figure 19. Each level of inference is also optimal for a particular type of signal change: Peak-level inferences are more sensitive for large signal amplitudes (regardless of signal extent), whereas cluster-level inferences are more sensitive for extensive signal change over a low height threshold. Set-level inferences are more sensitive for signal change across multiple clusters, regardless of signal height or extent. Because I hypothesise localisation of specific language functions when contrasting tasks to their respective sensory baseline condition, these findings shall be reported at the peak level. All group map analyses are described in the results section of each chapter, and significant peaks (>8mm apart) are provided in tables in the relevant section of the appendix. Tables include MNI co-ordinates, cluster size, *Z* values and approximate anatomic labels, based on the Automatic Anatomic Labels (Tzourio-Mazoyer et al., 2002) provided by xjview (http://www.alivelearn.net/xjview8/). Cerebellar activation is not reported for any of the analyses in the following chapters, as coverage of the cerebellum was variable across the sample and the explicit mask excluded voxels in the most variable (lower) sections.

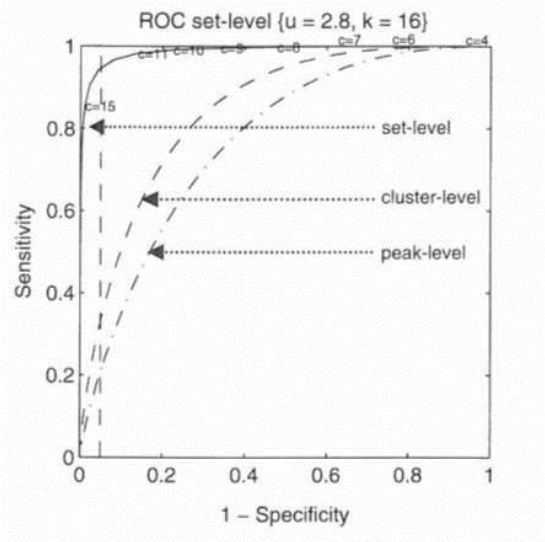


Figure 19: Sensitivity and specificity of inferences made at peak-, cluster- and set-level (from Friston 2007).

8. DEVELOPMENT OF A PRE-SCAN PREPARATION PROTOCOL

Very young children (< 7 years) show very high scan failure rates, of up to 57% (Bryars et al 2002). Scan success rates increase with age however (Byars et al., 2002; Yerys et al., 2009) as key executive functions develop, such as the ability to inhibit movement and sustain attention (Anderson, 2002). Many studies have begun to adapt their MR protocol to improve scan success rates in paediatric populations (Byars et al., 2002; de Bie et al., 2010; Gaillard et al., 2004; Hallowell, Stewart, de Amorim, & Ditchfield, 2008; Raschle et al., 2009; Schlund et al., 2011; Thomason, 2009). These studies highlight several variables associated with scan failure which can be addressed to improve the yield of fMRI in paediatric and patient populations (Table 11). Encouragingly, initial research suggests high scan failure rates in children can be reduced when appropriate pre-scan preparation protocols are put in place (de Bie et al., 2010; Hallowell et al., 2008). Design solutions shown to be effective at targeting common causes of scan failure in the literature (Table 11) were therefore implemented in a pre-scan preparation protocol for the Panda Games (Figure 20).

Study	Cause of scan failure	% of scan failures	Design solution
Croft et al., 2013	Poor task performance	60%	Pre-scan task practice Age-adjusted tasks (Gaillard et al., 2001) In-scanner performance monitoring (Croft et al., 2013)
Byars et al., 2002	Anxiety or distress	57%	Pre-scan information (Chesson et al., 2002) Child-friendly preperation (Raschle et al., 2009) Mock scanning (de Bie et al., 2010)
Yerys et al., 2009	In-scanner movement	16%	Pre-scan information; why it is important to keep still and how (Thomason, 2009) Mock scanning (de Bie et al., 2010)

Table 11: Causes and rates of scan failure in children, along with design solutions for a pre-scan preparation protocol aimed at addressing each causal factor.

8.1 A pre-scan protocol for children

8.1.1 At home

I created age-appropriate online 'How to play' booklets, which featured a cartoon character 'Super Panda' (Appendix8.). For children aged 5-7 years, the booklet opened with a picture story ('*Super-Panda's brain scan adventure*') which aimed to introduce very young children to the procedures they would be going through during participation in the research study (Pressdee, May, Eastman, & Grier, 1997). This story showed Super-Panda undergoing a brain scan. It showed participants where they would lie inside the scanner, referred to the noises the scanner made, and emphasised the importance of staying still inside the scanner. The *'Your brain scan'* page introduced participants to what a brain scanner looks like, told them they could bring their own favourite DVD to watch inside the scanner, described what fMRI is used for, and provided a click-on link to a video of someone undergoing fMRI at Great Ormond Street Hospital NHS Foundation Trust. A whole page was dedicated to explaining why it is important to keep still in the scanner, using picture examples (Thomason, 2009). Another page explained the importance of speaking softly inside the scanner, to help with keeping still. Following this preparatory section the Panda Games were introduced, with click on links to online practice versions of each task, and online examples of the Alien Game stimuli (showing the correct answer). A '*How the games work'* page showed what would happen for every task, step-by-step. Each task was then introduced in more detail, including; 1) a brief description of what would happen and what the participant had to do, 2) a picture example, 3) the reason the participant had to perform this task ('*What is this for?*'), 4) task instructions ('*Game rules'*), and 5) answers to frequently asked questions. The age-appropriate version of this booklet was sent to participants before their appointment.

8.1.2 On the day (prior to scanning)

8.1.2.1 Picture wall

Participants were shown a picture wall, which included photos of the radiographer ("*the person who takes the photos*"), the scanner and its parts, and

the type of images obtained from MRI and fMRI. These were used as discussion points to gauge the child's current understanding of the procedure, to answer any questions about what was going to happen, and to explain certain aspects of MR scanning. For example, the head coil was introduced as a space helmet, which works like a zoom on a camera ("*it helps the scanner zoom in on your brain, and get a better photo*").

8.1.2.2 Task practice

Participants were guided through the task instructions using hard-copies of the 'How to play' booklet, to gauge their understanding of task instructions. Additional practice was performed at this point using a computer, if required.

8.1.2.3 Mock scanning

Very young participants (aged 5-7 years) or participants who were anxious underwent mock scanning. Mock scanning was not always possible due to limitations regarding appointment times. The mock scanner was built from a children's play tunnel, pillow and head coil. Depending on age and size, children lied in or out of the tunnel - with the head coil closed - and listened to scanner sounds at increasing volumes.

8.1.3 During scanning

Inside the scanner children were able to speak with the radiographer and myself via an intercom system and MR compatible microphone. Participants were given verbal encouragement and reward throughout the scanning session. Further, several design aspects of the language tasks used during scanning were optimised to avoid scan failure:

8.1.3.1 Short language tasks

A blocked event-related design (section 9.2) was used to optimise fMRI, such that shorter task blocks (4 minutes) could be used, to shorten the duration of sustained attention-to-task and to reduce overall scanning time.

8.1.3.2 Four difficulty levels

Each task had four difficulty levels; Very easy (for children aged 5-7 years), easy (8-10 years), moderate (11-13 years), and difficult (14-16 years).

Difficulty was determined by name agreement, response uncertainty, and age of acquisition for the target item (see Chapter 2).

8.1.3.3 Overt speech response mode

Participants responded for each task using overt speech, allowing online performance monitoring (Chapter 1). Poor performance could be identified early in the scanning session, allowing for task instructions to be repeated, or for the task to be aborted and started again.

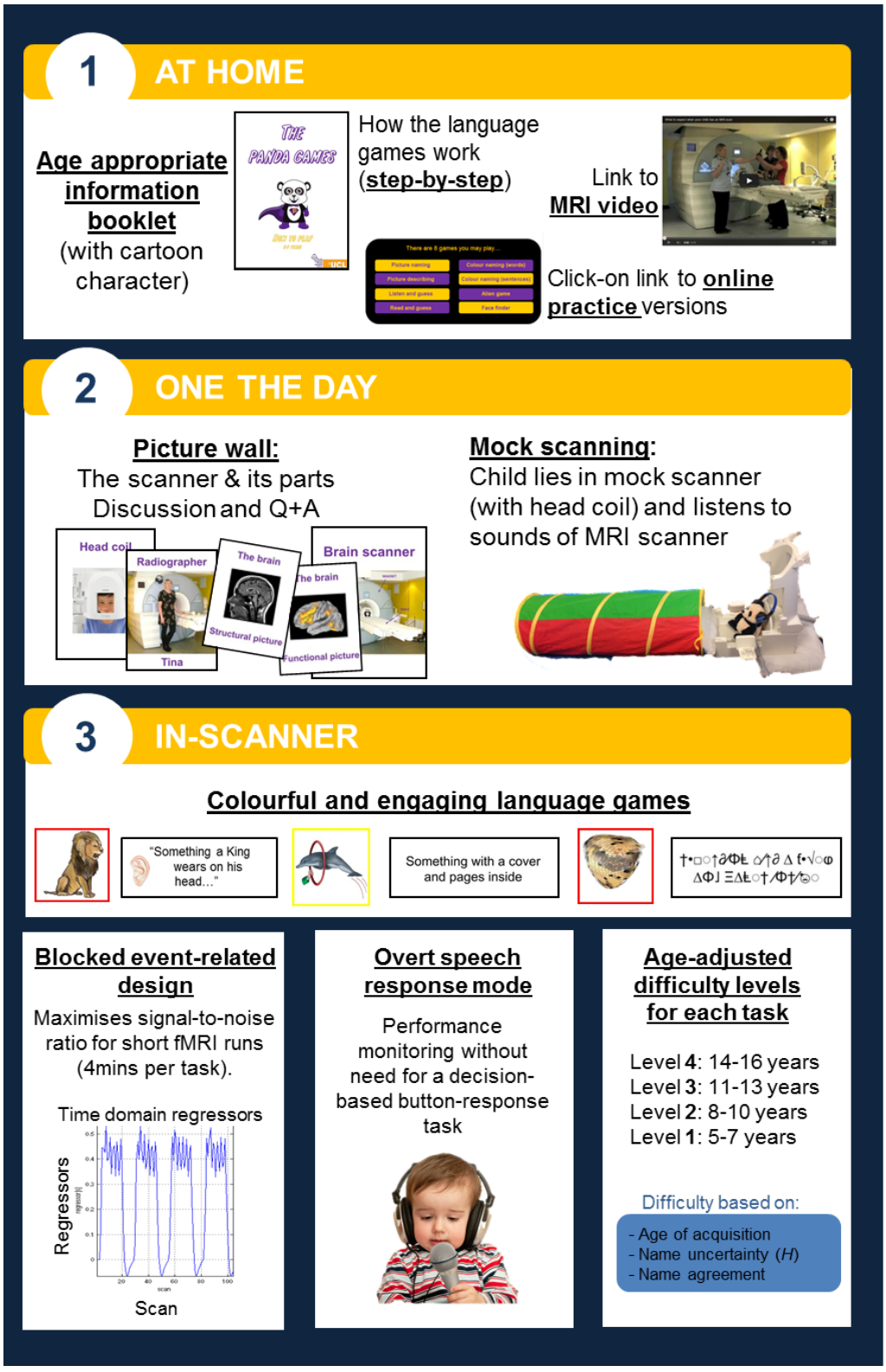


Figure 20: A pre-scan preparation protocol for children. Each stage has been designed to target common causes of scan failure, including anxiety, poor task performance and in-scanner movement.

9. THE PANDA GAMES FMRI TASK BATTERY

For each task, twenty four stimuli were presented in a blocked event related design. There were four task blocks separated by 16 seconds of rest (with fixation to cross hairs). Within each task block six stimuli were presented every six seconds, with a total task block duration of 36 seconds. To prepare the child to look/listen, and to prevent a startle response in response to the first auditory stimulus in a block, task blocks were preceded by a 'get ready' stimulus (picture of an amber traffic light, presented for 2.16 seconds). Each task lasted a total duration of three minutes and 49 seconds. Tasks were performed in pairs (task and baseline). The order within each pair and between pairs was counterbalanced across the sample using a Latin Square design. The entire Panda Games task battery had a total duration of 30 minutes. A shorter version of the Panda Games task battery, which excluded *Read and Name* and its sensory baseline task (*Face Finder*), had a total duration of 23 minutes and was performed by young children (aged 5-7 years) and children with epilepsy.

9.1 Task design

For every task, 24 stimuli were selected for each difficulty level (Chapter 2). Stimuli descriptives (Chapter 2, section 4.2), examples (Appendix 1.) and task instructions are provided (Appendix 8.). For each task, stimulus length and the distribution of living/non-living and natural/man-made items were well-matched across difficulty levels (Chapter 2, Table 6- Table 9).

9.1.1 Picture Naming (PN)

This task was designed to induce visual processing, word retrieval from visual input, and articulation. Participants were asked to name 24 colour pictures depicting objects and animals. Participants were asked to name items as quickly as they could, and to say the first thing that "popped into their head". Images were 350x350 pixels, surrounded by a 5 pixel border in red, yellow, green or blue. Items were assigned to a border colour using a random sequence generator, such that a quarter of items had a border of each colour, in each difficulty level. Border colours were consistent between stimuli in the Picture Naming condition and the baseline (Colour Naming with words) stimuli.

9.1.2 Colour Naming (with words) (CNw)

This task was designed to induce visual processing, semantically-shallow word retrieval from visual input, and articulation. Participants were asked to name the colour of the square border (5 pixels) around a 350x350 pixel meaningless coloured pattern. For each difficulty level, 25% of stimuli had square borders that were red, 25% that were yellow, 25% that were green and 25% that were blue, as in the Picture Naming condition. The meaningless pattern was a visually warped version of Picture Naming stimuli, matched for visual complexity. Stimulus order matched the Picture Naming task. Each stimulus was presented for four seconds, followed by a two-second cross hair. By contrasting PN with CNw, I aimed to localise processing associated with semantic processing during word retrieval.

9.1.3 Picture Describing (PD)

This task was designed to induce visual processing, sentence-level semantic and syntactic processing from visual input, and articulation. Participants were asked to describe what was happening in the picture, using one simple sentence with a subject-verb-object (S-V-O) structure. Each 350x350 pixel colour picture showed a subject and an object, with the subject involved in one of four actions; eating, drinking, jumping or falling. Participants were asked to respond in the active present tense ("as if it's happening now"), e.g. "The squirrel is jumping from the tree", as opposed to "The squirrel jumped from the tree". For drinking stimuli, participants were asked to say what the subject was drinking from, as opposed to what the subject was drinking. The distribution of verbs was well-matched between difficulty levels (see Chapter 2, Table 9). Each picture was surrounded by a 5 pixel colour border, as for Picture Naming. Each stimulus was presented for four seconds, followed by a two-second cross hair.

9.1.4 Colour Naming (with sentences) (CNs)

This task was designed to induce visual processing, sentence repetition with semantically shallow word retrieval from visual input, and articulation. Participants were asked to name the colour of the square border (5 pixels) around a 350x350 pixel meaningless coloured pattern, using the same sentence structure each time; "The colour of the square is…". Participants were asked to say the

correct colour for each stimulus at the end of the sentence structure. For each difficulty level, 25% of stimuli had square borders that were red, 25% that were yellow, 25% that were green and 25% that were blue, as in the Picture Describing condition. The meaningless pattern was a visually warped version of Picture Describing stimuli, matched for visual complexity. Stimulus order matched the Picture Describing task. Each stimulus was presented for four seconds, followed by a two-second cross hair. By contrasting PD with CNs, I aimed to localise processing associated with sentence-level semantic and syntactic processing.

9.1.5 Listen and Name (LN)

This task was designed to induce complex auditory processing, sentence level semantic and syntactic processing from phonological input, word retrieval and articulation. Participants were asked to listen to sentence-level auditory descriptions of objects and animals, spoken by a male (50% of trials) or female (50% trials), with a mean length of 7 words and mean duration of 2.39 seconds (range=1-5 seconds). Sentences had been recorded by a male and a female speaker of the same age (24 years), in a sound proof room using prorec software (http://www.phon.ucl.ac.uk/resource/prorec/). Root mean square (rms) amplitude was normalised across both speakers, to produce stimuli with equal average intensity. Sentence descriptions were taken from child versions of the Oxford English Dictionary: Oxford First Dictionary (5-7 years), Oxford Primary Dictionary (7-11 years), or the Oxford Dictionary For Schools (11-16 years) (described further in Chapter 2, section 4.2.3). The four difficulty levels were matched for stimulus length (see Chapter 2, Table 8), which affects the extent of activation in superior temporal regions (Yeatman et al., 2010).

9.1.6 Alien Game (AG)

This task was designed to induce complex auditory processing, semantically shallow word retrieval, and articulation. Participants were asked to listen to 'alien's' talking, and to name the sex of each alien ("Is it a boy or a girl?"). Alien speech stimuli were created by spectrally rotating (inverting) LN stimuli using MATLAB (scripts available here: http://www.phon.ucl.ac.uk/resource/software.php) as reported previously (Scott,

Blank, Rosen, & Wise, 2000). This script performed spectral rotation (around 2 kilohertz; kHz) using a digital version of the simple modulation technique (Blesser, 1972). Stimuli were then equalized with a high-pass filter, so that the rotated and original stimuli had approximately the same long-term spectrum, followed by low-pass filtering at 3.8 kHz. Rotated speech provides a high-level control for the LN task, as stimuli are unintelligible but matched for duration, and temporal and spectral complexity, including; intonation, pitch, and some phonetic features (Blesser, 1972). Participants were able to make the gender decision using pitch information. By contrasting LN with AG I aimed to localise sentence-level semantic and syntactic processing from phonological input.

9.1.7 Read and Name (RN)

This task was designed to induce visual processing, sentence level semantic and syntactic processing from orthographical input, word retrieval and articulation. Participants were asked to read sentence-level descriptions of objects and animals covertly ("In your head"), and to say the name of the item being described out loud; similar to LN. Sentence descriptions were taken from child versions of the Oxford English Dictionary: Oxford First Dictionary (5-7 years), Oxford Primary Dictionary (7-11 years), or the Oxford Dictionary For Schools (11-16 years) (described further in Chapter 2, section 4.2.3). Mean stimulus length was 7 words. Stimuli were presented on the screen for five seconds, followed by a cross-hair for one second. Stimuli were created using Arial font (size 44) in Microsoft Powerpoint, and were saved as jpeg files.

9.1.8 Face Finder (FF)

This task was designed to induce visual processing (including left to right horizontal search; similar to eye movements during reading), semantically shallow word retrieval, and articulation. Participants were asked to look along a line of symbols (from left to right; "like if you're reading") to find the smiley face icon, at which point they had to say if the face was a girl or boy (as for AG). To ensure no names or phonological representations were automatically activated by written stimuli, I used unfamiliar abstract symbols as opposed to nameable symbols (e.g. digits) or pseudowords. To create the stimuli I assigned a symbol to

each letter of the alphabet, and then replaced letters in the RN stimuli with their corresponding symbol (maintaining the spaces between words). For each stimulus, one symbol was replaced with a face icon. The face icon could take one of 7 positions in the stimulus, replacing a symbol located between 65-95% of the way through the symbol string (at 5% intervals). By positioning the face within the latter half of the symbol string, participants had to scan along the length of the stimulus from left to right (similar to eye movements during reading) and responses were made toward the end of the stimuli, consistent with RN. Two face icons - a male and female - were created using Microsoft Powerpoint and were saved as jpeg files. Half of stimuli included the male icon, half included the female icon. Previous studies used less visually complex stimuli for a reading baseline task, such as black and white dots (Berl et al., 2010) or lines of differing orientation (Constable et al., 2004). However, identification of the face icon was immediate and required less visual attention when placed among such visually simple stimuli. Among the abstract symbols used here, identification of the face icon required more attention to the stimuli. Symbols also provided a better match for visual complexity versus simple visual stimuli. By contrasting RN with FF I aimed to localise sentence-level semantic and syntactic processing from orthographical input.

9.2 First-level design matrix

The onset and duration of each stimulus was modelled using an event-related design. Event-related analyses of blocked stimuli provides more accurate modelling of the HRF in language tasks, and reduces residual error (Mechelli, Henson, Price, & Friston, 2003; Price, Veltman, Ashburner, Josephs, & Friston, 1999). Event-related designs are also more sensitive for detecting language cortex in the pre-surgical setting compared to blocked designs (Tie et al., 2009). Further, event-related analyses allow trials to be modelled according to response type (correct/incorrect) (Figure 21). The actual stimulus duration was modelled for PD (4 seconds), LN (mean=2.39 seconds), RN (5 seconds), and their respective baseline conditions. Picture Naming was modelled with an event duration of 0 seconds. Each language task was modelled in its own design matrix (session 1), along with its respective

baseline condition (session 2). Three stimulus functions were used to model three conditions for each task: 1) correct response trials, 2) incorrect response/omission trials, and 3) 'get ready' stimuli (Figure 21). Temporal and dispersion derivatives were also included in the design matrix for each stimulus function, to reduce the residual error associated with between-subject variability in the onset or duration of the BOLD response (Figure 21). This also ensured age-related changes in network activation did not reflect variance in the onset or dispersion of the BOLD signal which may be related to development of neurovascular coupling (see section 5.1.4). Realignment parameters were entered into the design matrix for each task as effects of no interest. Data were high-pass filtered such that the effects of slow signal drifts lasting longer than 128 seconds were removed. Second-level analyses are described in each chapter. Beta weights were estimated for all voxels (regardless of intensity values) within a grey matter mask to reduce the number of comparisons performed. This also prevented exclusion of low-signal voxels from calculation of the beta weights, which can have serious consequences at the second-level (areas excluded during individual analyses at the first level are also excluded from group-level analyses, regardless of activation in these regions in other participants).

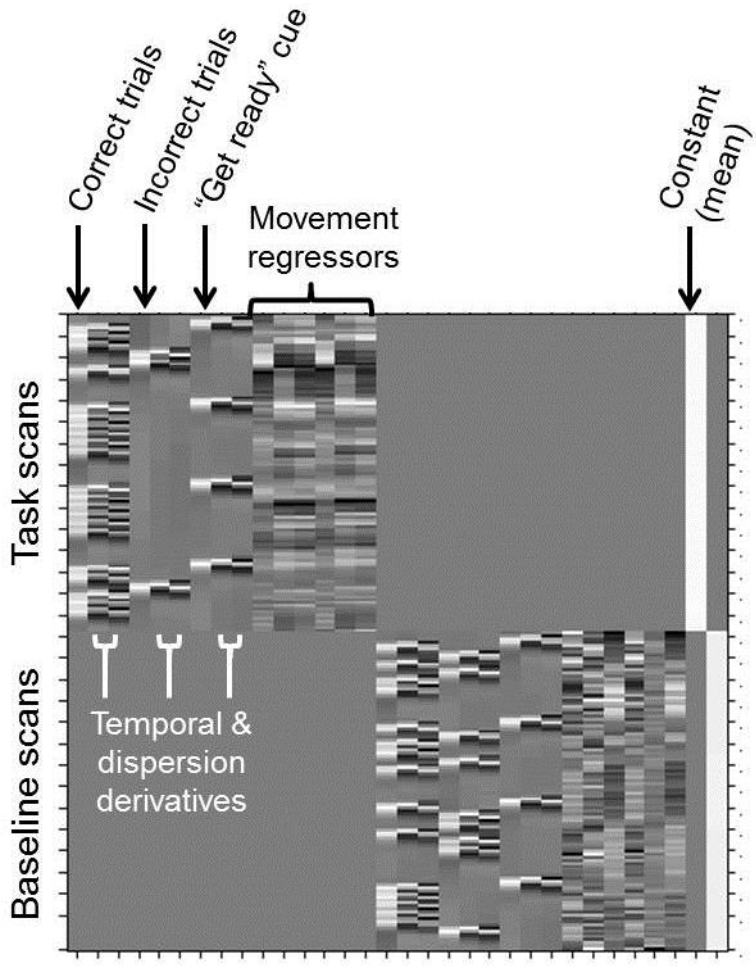


Figure 21: First-level design matrix for the Panda Games.

Chapter 4: Identification and repair of speech related movement artefacts in children

1. ABSTRACT

In this chapter I aimed to identify the optimal post-processing method for repairing speech-related movement artefact in fMRI data, to be used on data reported in the remainder of this thesis. Forty three healthy children performed the Panda Games during fMRI scanning. I rated the severity of artefact in a low-movement condition (Listen and Name) and a high-movement condition (Picture Describing). I then compared three post-processing artefact repair methods (scan nulling, robust weighted least squares, and motion fingerprint) with the standard preprocessing pipeline, to identify which method produced the highest quality data, as assessed by subjective and objective measures. Subjective artefact ratings showed no difference in data quality between repair methods. An objective measure of data quality showed a significant reduction in data quality following repair using robust WLS, due to reduced sensitivity. Although no single method improved the specific form of movement artefact observed in this dataset (which occurred within the volume, rather than between volumes), the motion finger print method aims to account for the widest range of movement-induced artefacts, while maximising sensitivity (by minimising the number of regressors added to the GLM). Further, there was a trend for increased data quality using the motion fingerprint method in participants with the most severe artefacts in the high-movement condition. I therefore conclude motion finger print is the optimal post-processing method for data reported in the remainder of this thesis.

2. AIMS

In this chapter I introduce three post-processing methods which have been developed to reduce the effect of several types of movement-related artefacts on scan quality. In the first of two experiments I then summarise the prevalence and severity of speech-related movement artefacts in my cohort, and the effects of task, age and sex on artefact severity. In the second experiment I compare the effectiveness of the three post-processing methods outlined in the introduction, which include: 1) a volterra expansion of realignment parameters plus scan nulling (Lemieux et al., 2007), 2) robust weighted least squares (Diedrichsen & Shadmehr, 2005), and 3) the motion fingerprint (Wilke, 2012). I compare the effectiveness of these methods to repair sub-TR speech-related movement artefacts, assessed using subjective artefact ratings. I also compare the effectiveness of each method to detect activation in target language networks, using an objective measure of scan quality. To conclude, I select the optimal method for use in the remaining chapters of this thesis, which focus on investigations of the developing language system and require stringent data quality control.

3. INTRODUCTION

Here I introduce three post-processing methods, which have been developed to address a range of speech-related movement artefacts once data has been obtained. These post-processing methods each target different movement-related artefacts, explained in the following sections.

3.1 Post-processing methods to repair speech-related movement artefacts

3.1.1 Movement regressors in the standard GLM (Friston et al., 1996)

As part of the standard pre-processing pipeline, images are realigned to the first scan in a session. This produces a series of six realignment parameters, which describe translations (in *x*,*y* and *z* directions) and rotations (pitch, roll and yaw) that occur throughout the scanning session. These movements cause voxels to be displaced over time. Realignment aims to correct for this, and tries to ensure all voxels are aligned over time. In the standard GLM, these realignment parameters are included as covariates of no interest, to account for signal changes associated

with movement of the head from scan to scan. Typically, participants are excluded from group-level analyses if the maximum movement in any of the six directions exceeds a certain threshold. In adults this threshold is usually 1mm. In children and patients however, this threshold is often relaxed to 1 voxel (~3mm), due to increased in-scanner movement in these populations (Yerys et al., 2009). This method does not correct for several types of speech-related movement artefact, which are discussed below and are addressed by more advanced motion correction methods.

3.1.2 Volterra expansion and scan nulling (Lemieux et al., 2007)

This method differs from the standard GLM in two ways. First, 24 realignment parameters are included in the GLM (Figure 22). These are produced by a volterra expansion of the standard realignment parameters, which aim to account for spin excitation history effects in each of the six standard directions (*x*,*y*,*z*, pitch, roll and yaw). This method has been shown to be effective at explaining even small (<0.5mm) motion-induced artefacts across the brain (Lemieux et al., 2007). Further, it explains more variance than the six realignment parameters in the standard GLM (Friston et al., 1996; Lemieux et al., 2007) and the motion fingerprint method (Wilke, 2012), which is discussed below. The second way in which this method differs from the standard GLM, is the inclusion of scan nulling regressors (Figure 22) which account for variation in the BOLD response associated with large movements (over a specified threshold of mm/TR). By accounting for large movements as well as more subtle spin-excitation history effects, this method aims to reduce the residuals of the GLM, and thus increase the sensitivity to detect task-related activation. It represents a trade-off between discarding data entirely (as with ArtRepair in Chapter 1) and the standard GLM, which fails to adequately account for movement artefacts and therefore suffers from a low signal-to-noise ratio. However, unfortunately for this method, sensitivity to detect task-related activation decreases as a function of increasing the number of covariates in the GLM; at a rate of ~1% per additional regressor (Wilke, 2012). In an event-related design, sensitivity can decrease by as much as 9.51% when just the standard six covariates are added to the GLM, and by as much as 43.76% when 24 covariates are included (i.e. the expanded realignment

parameters) (Wilke, 2012). This reduction in detection power is carried through to the second-level, with decreases in sensitivity as much as 33.25%. As the method with the most regressors, I therefore hypothesise this expanded realignment parameters and scan nulling method (from here on referred to just as 'scan nulling') will produce poorer quality data relative to the other methods presented here, when quality is assessed according to sensitivity.

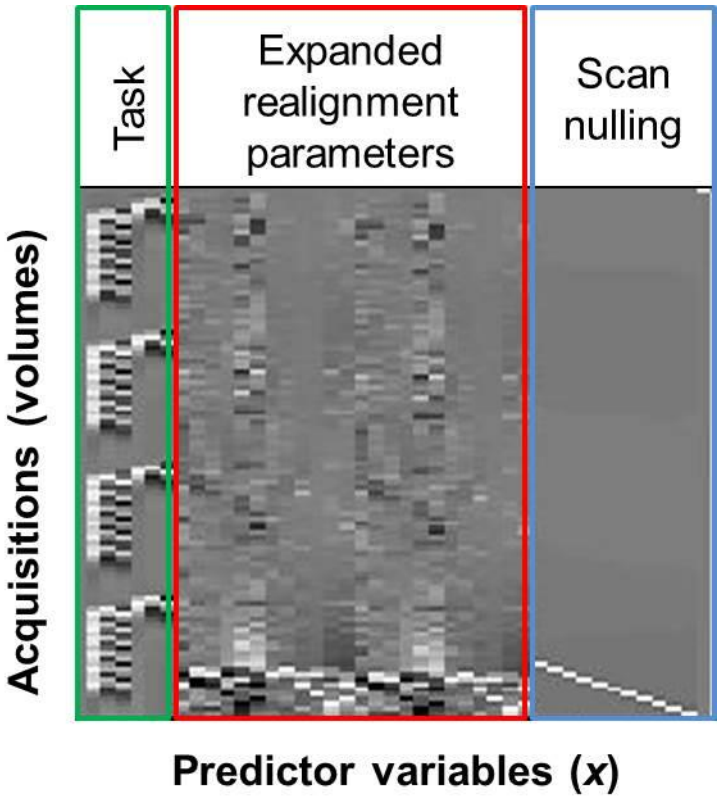


Figure 22: First-level (within-subject) design matrix including expanded realignment parameters (red) to account for variance associated with spin-excitation history effects, and scan nulling regressors (blue) to account for large movements (i.e. head jerks), in addition to the standard task-related regressors (green).

3.1.3 Robust weighted least squares (Diedrichsen & Shadmehr, 2005)

Essentially, this method uses the residuals from the GLM to de-weight noisy volumes during calculation of the beta weights. When the GLM is calculated, it produces the beta weights (β) for each predictor variable (x) as well as a measure of error (the residuals; e). The residuals reflect variance in the fMRI data (y) not accounted for by the linear combination of predictor variables and their beta weights ($x\beta_1 + x\beta_2 + x\beta_3\ldots$). The residuals are summarised in the residual mean squares (ResMS) image, which provides a measure of error for each voxel in the brain collapsed across time points. In this image brighter voxels represent higher values, and therefore highlight parts of the brain where the BOLD response (y) is less well explained by the predictor variables (Figure 23). In order to perform WLS, the residuals are summarised differently. Instead of being summed across the time series, residuals are summed across the volume. This results in a single measure of error for the entire brain, at each time point (Figure 24).

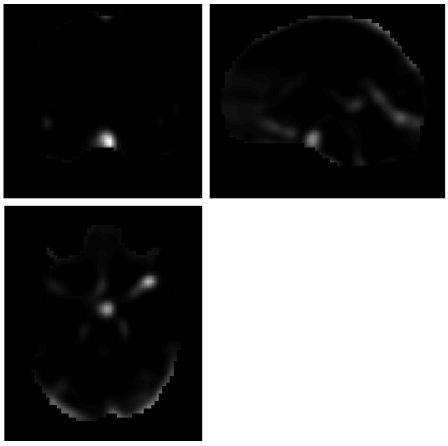


Figure 23: Residual mean squares (ResMS) image, produced by the standard GLM, displayed in three planes: coronal (top left), sagittal (top right) and axial (bottom left). Lighter voxels indicate higher variance in residuals (more error); brain regions

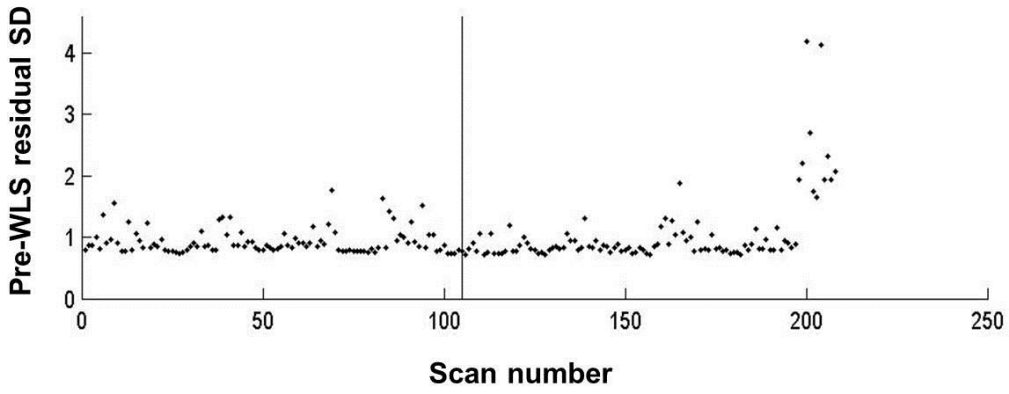


Figure 24: Standard deviation of residuals for each time point throughout the timeseries, calculated prior to performing robust WLS. Data are shown for the youngest participant in this cohort, for the Listen and Name Game (scan 0-104) and Alien Game (scans 105-209; onset marked by the vertical line). Volumes toward the end of the Alien Game show large residuals.

Because the residuals are summed across each volume, it is important to account for non-sphericity (non-uniform variance in the differences in error from TR to TR; between volumes). Non-sphericity can be accounted for by using a restricted maximum likelihood (ReML) approach to estimate beta weights and the residuals. The unbiased estimates of residuals are then inverted, and used to weight volumes during calculation of least squares (WLS). This leads to what the authors refer to as a "soft" exclusion of images; noisy volumes are de-weighted so they contribute less during calculation of the beta weights (instead of being effectively excluded from the calculation altogether, as with scan nulling). As a data-driven approach, robust WLS also avoids the need for a subjective threshold to identify noisy volumes, as in scan nulling. This method is designed to address unique discrete motion events which occur during acquisition of a volume. It is therefore hypothesised to provide the best repair for sub-TR movement artefacts. However, whether de-weighting noisy volumes has an effect on sensitivity to detect activation in the context of task-correlated motion remains to be seen. Because speech movements are correlated with the task in the Panda Games, I hypothesise this method will result in reduced sensitivity to detect activation in language networks.

3.1.4 Motion fingerprint (Wilke et al., 2012)

The advantage of this method relative to the previous methods, is that as well as accounting for scan-to-scan head movement and spin-excitation history effects, it also accounts for reslicing interpolation errors and the interaction of movement with individual brain shape and position in the scanner (explained further below). Further, it achieves all this whilst keeping the number of regressors in the GLM minimal compared to the expanded regressors and scan nulling method, which – as we have seen in section 3.1.2 - is an advantage in terms of sensitivity to detect task-related activation (Wilke, 2012).

In total this method produces six regressors for the first-level GLM. Three of these are synthetic time courses. The whole-brain synthetic time series is based on the first scan from a session replicated *n* times (where *n* is the total number of scans in a session), such that signal change over time is 0. The inverse of the realignment parameters are then applied to this time series, to introduce the original displacement of head position over time, and any interpolation errors introduced by the realignment procedure. The whole-brain synthetic time series therefore reflects the expected signal change over time, associated purely with movement. The motion fingerprint method should be more sensitive for detecting task-related activation in the presence of task-correlated motion (modelled by the synthetic timeseries), compared to methods which de-weight or null task-related time points with high motion (i.e. scan nulling and robust WLS).

The motion fingerprint method also accounts for the interaction of head rotations with B_0 (Andersson, Ashburner, & Friston, 2001; Hutton et al., 2002), whereby the displacement of a voxel is a function of its distance from the origin of the volume (i.e. the centre of the gradient coil) (Figure 25A). To account for these regional motion effects, the mean total displacement is calculated for each region of grey matter, according to a reference brain with 116 cortical and subcortical parcellations (Tzourio-Mazoyer et al., 2002), and voxels are labelled according to their maximum total displacement over time (Figure 25B). This information is

used to adjust the synthetic timeseries. Because the motion finger print models these regional motion effects, this method should account for more within-subject signal variance, relative to the other methods described here.

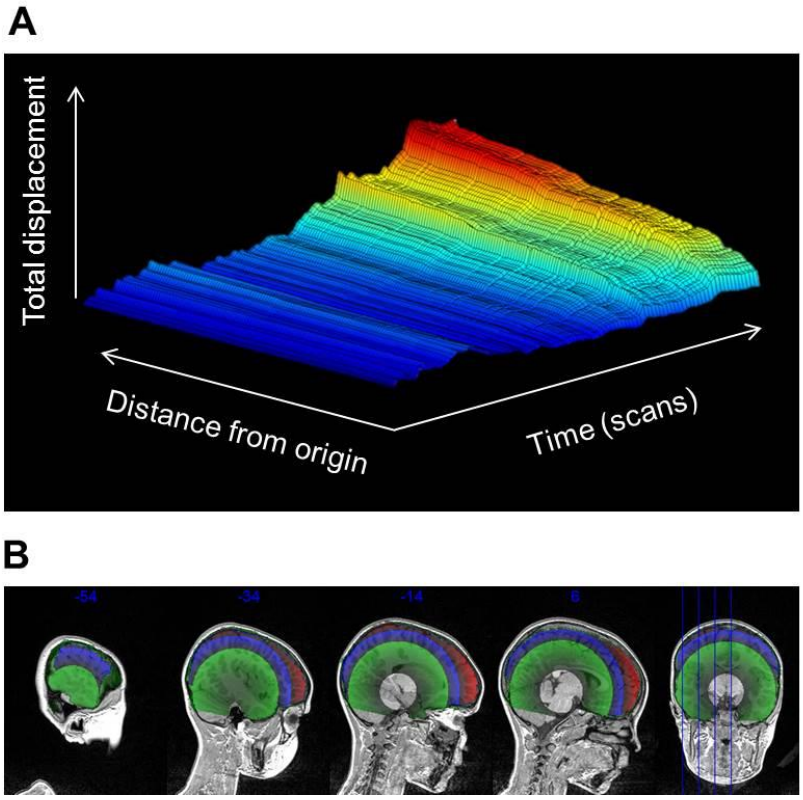


Figure 25: A) Adaptation of a graph taken from Wilke (2012), showing total displacement of a voxel (Y axis) as a function of brain region (X axis) over time (Z axis) in a single participant. Displacement is greatest for regions far from the origin of the volume towards the end of the scan session. B) Example of an individual motion mask for the youngest participant in this sample (aged 5) during the Listen and Name game, created by the motion fingerprint method and displayed across sagittal slices. Colour code shows brain regions with movement >1mm (green), > 2mm (blue) and >3mm (red), in relation to the origin of the volume (central sphere). Slices are depicted on the far right. This information is used to adapt the synthetic timeseries to account for motion*B_0 artefacts in the motion fingerprint method.

Seed points are placed in each of the 8 corners of the whole-brain synthetic volume. These are moved inwards on a voxelwise basis until they reach the edge of the brain. At this point, a 10 voxel diameter cube ROI is drawn around each seed. In addition, a ninth cubic ROI is drawn around the centre of mass of the brain volume. Mean time courses are then extracted from these ROIs (based only on voxels with signal change above the mean; in order to exclude the influence of background voxels which are partly included in the 8 corner ROIs). Including just three of these regional time courses in the GLM has been shown to explain as much variance as the standard six realignment parameters (Wilke 2012). As such, the three regional time courses which share the least variance are combined to form the 'motion fingerprint'; a unique description of movement for a given subject during a given scan session. In order to account for the spin-excitation history effects, three additional versions of the synthetic timeseries regressors are created, by shifting each one by a single time-point. These time shifted regressors are included as part of the motion fingerprint.

3.2 Artefact severity and repair in the Panda Games cohort

In the following two experiments I determine the prevalence and severity of speech-related movement artefact, and identify risk factors for severe movement artefact (Experiment 1). I then compare the four methods outlined above in terms of their ability to a) repair sub-TR movement artefacts, and b) improve scan quality (Experiment 2). I perform these analyses in order to identify the optimal GLM for producing high quality group maps in future chapters, with stringent control of both the type I and type II error rates.

4. EXPERIMENT 1: ARTEFACT PREVALENCE

In this section I aim to characterise the severity of speech-related movement artefacts in my cohort of healthy children and adolescents, identifying participant characteristics associated with poor data quality. I therefore have the following research questions and hypotheses:

1. What is the prevalence of speech-related movement artefact in this cohort of children and adolescents?

   Hypothesis: Movement related artefacts will be present in a large majority of data due to the use of an overt speech response mode (combined with an ascending slice acquisition method) and due to prolonged scanning times.

2. Which tasks are associated with the most speech-movement related artefacts?

   Hypothesis: Movement related artefacts will be increased in sentence-level production tasks relative to single word production tasks.

3. Which participant characteristics are associated with more severe artefacts?

   Hypothesis: Artefacts will be higher in younger children (aged 5-8 years) and males, due to evidence for increased in-scanner movement in these participants (Yerys et al., 2009).

5. METHODS

5.1 Sample

Because I wanted to capture the frequency and severity of movement-related artefacts in a representative developmental sample, all participants (N=43, 5-16 years, 49% female) were included regardless of how much they moved during scanning (which in some cases well exceeded the 1 voxel cut-off used in developmental studies).

5.2 Language fMRI

All four language tasks from the Panda Games task battery were included: Picture Naming (PN), Picture Describing (PD), Listen and Name (LN) and Read and Name (RN). Data were pre-processed as described in Chapter 3. Results for each task are presented for each task compared to its sensory-baseline task, unless otherwise stated.

5.3 Artefact ratings

I rated the severity of speech-related movement artefact by looking at the average signal change across each scan session (which included both the task and baseline). This contrast shows reductions in signal within the axial plane, associated with movements in the z direction and pitch rotations during speech production (Figure 26). A rating of 0 (no artefact) was given when voxel intensity was uniform across the volume (Figure 26, A). A rating of 1 (minor artefact) was given when a (thin) single line was present across the brain (Figure 26 B), or when two broken lines were apparent (Figure 26 C). A rating of 2 (severe artefact) was given when multiple lines (>2) were seen throughout the brain (Figure 26, D), or in the presence of thick axial bands of signal dropout (Figure 26, E). Ratings were performed when data had been modelled using the standard GLM, and also when data had been modelled with each of the three artefact repair methods. A second expert reviewer performed ratings using the same criteria, but blind to task, participant and repair method. The blind reviewer rated artefacts before repair (standard GLM) and after repair (robust WLS) for PD. This task was chosen as it was hypothesised to show the largest variance in artefact. Inter-rater agreement analysis was then performed to validate the consistency of my own (non-blind) ratings.

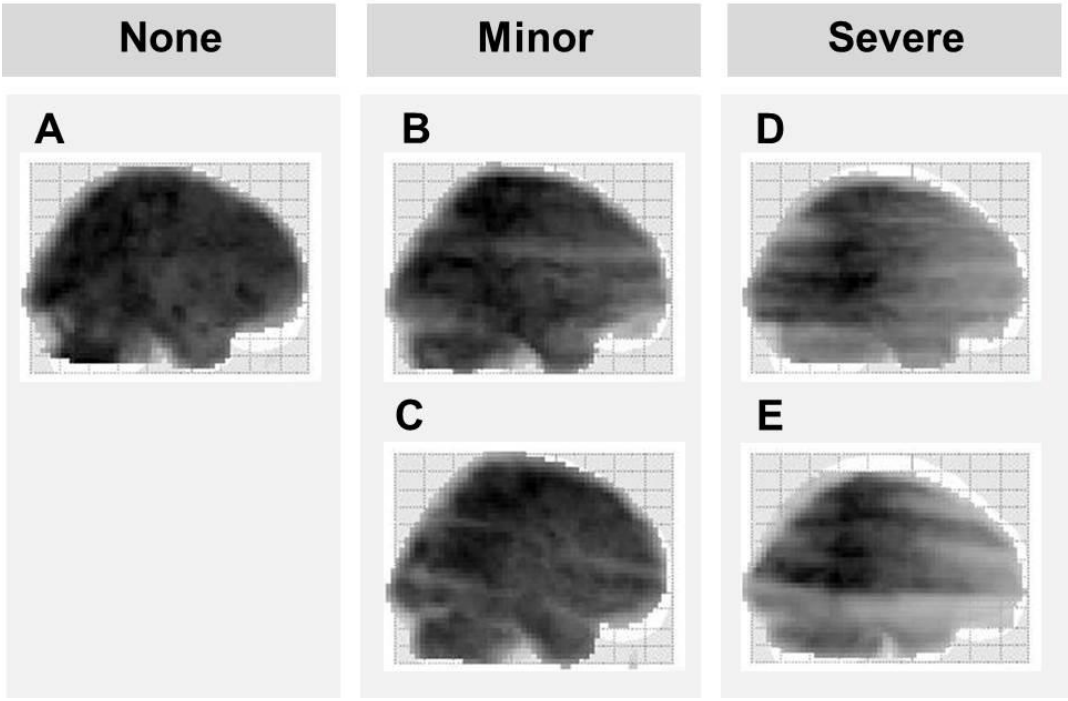


Figure 26: Examples of artefact free data (A) and data with speech-related movement artefact in the axial plane, rated as minor (B-C) or severe (D-E).

6. RESULTS

6.1 What is the prevalence of speech-related movement artefact in this cohort?

Inter-rater reliability analysis showed good agreement between blind and non-blind artefact ratings before artefact repair (average measures intra-class correlation coefficient = 0.88, $p<0.001$) and after artefact repair (average measures intra-class correlation coefficient = 0.7, $p<0.001$). According to subjective ratings, speech-related artefacts were identified in ~40% of data, across all tasks (Table 12).

	Artefact rating		
Task	**None**	**Mild**	**Severe**
PN	58%	33%	9%
PD	44%	28%	28%
LN	53%	33%	14%
RN	69%	31%	15%

Table 12: Percentage of participants showing mild and severe artefacts in the z plane due to sub-TR speech-related movement, for the total sample. PN (Picture Naming), PD (Picture Describing), LN (Listen and Name), RN (Read and Name).

6.2 Which tasks are associated with more severe artefacts?

Artefact ratings were compared between tasks using Friedman's ANOVA (a non-parametric repeated measures ANOVA for ordinal data), which showed a significant difference ($\chi^2(2)=6.72$, $p=0.04$). Post-hoc comparisons using the Wilcoxon signed ranks test (a non-parametric related samples t-test for ordinal data) with a Bonferroni correction for multiple comparisons ($\alpha=0.008$), found a trend for differences between PD and both PN ($p=0.03$) and LN ($p=0.06$), suggesting sentence production is associated with more severe artefacts than word production, as hypothesised.

6.3 Which participant characteristics are associated with more severe artefacts?

Artefact ratings are shown by age and by task in Figure 27. A Kruskal-Wallis test (a non-parametric one-way ANOVA for ordinal data) showed artefact severity differed by age group for both visual tasks: PN ($H(2)=7.12$, $p=0.03$) and PD ($H(2)=8.78$, $p=0.01$). There was also a trend for age effects in both auditory tasks: LN ($p=0.08$) and RN ($p=0.07$). Two post-hoc Mann-Whitney U tests were performed with Bonferroni correction ($\alpha=0.01$) to compare artefact ratings for PN and PD between age groups (5-8 versus 9-12 years, and 9-12 versus 13-16 years). The youngest age group (5-8 years) had more severe artefacts compared to children aged 9-12 years for both PN ($U=56$, $p=0.01$) and PD ($U=40.50$, $p=0.001$). Children aged 9-12 years did not differ from children aged 13-16 years, although there was a slight trend for more artefact in 9-12 year olds for PD (one-tailed $p=0.04$). There were no sex differences in artefact severity (Mann-Whitney U test, all p values >0.13).

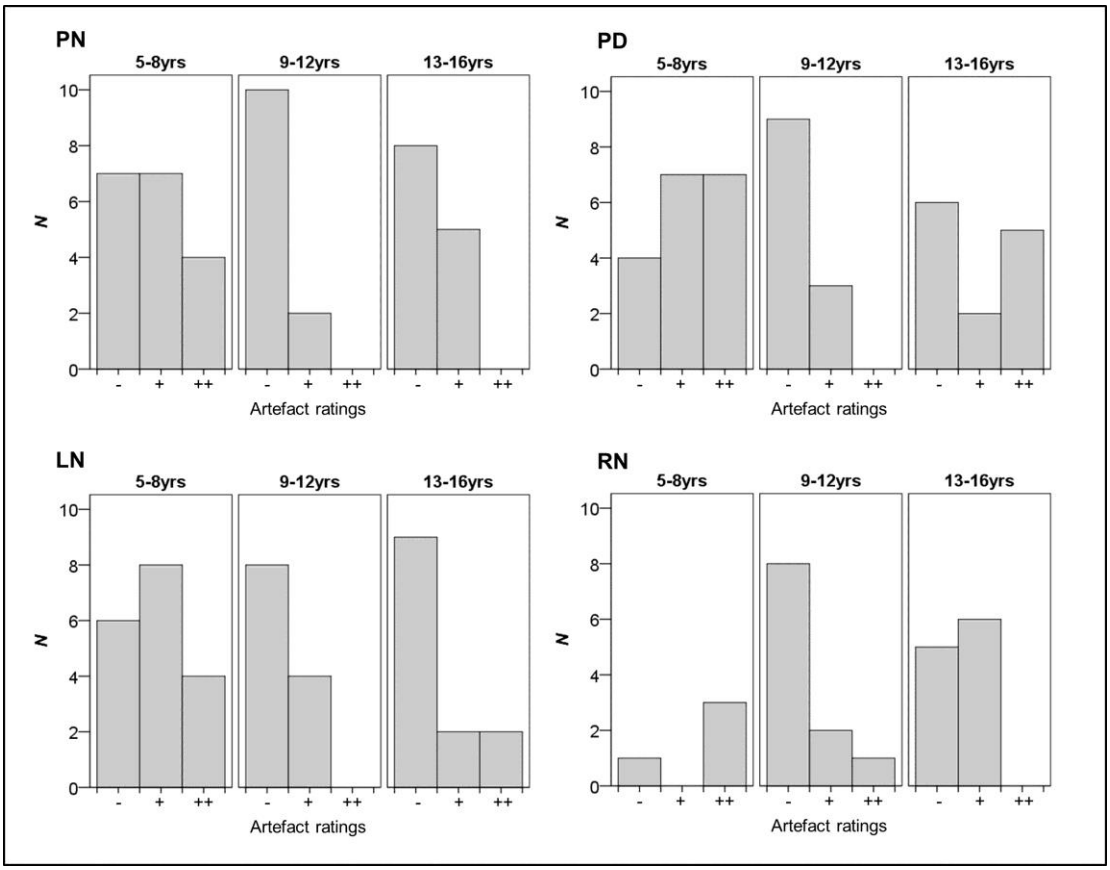


Figure 27: The frequency of acquiring data without artefact (-), and with mild (+) or severe (++) speech-related movement artefacts in children and adolescents, shown for all language tasks in the Panda Games battery: PN (Picture Naming), PD (Picture Describing),

7. EXPERIMENT 2: EFFECT OF POST-PROCESSING REPAIR METHODS

Here I compare three post-processing methods in terms of their effectiveness for reducing artefact and increasing data quality (as assessed by sensitivity and specificity to detect activation in language networks). I hypothesise robust WLS or scan nulling will show the greatest improvement in sub-TR movement artefact (as assessed by subjective ratings) due to de-weighting or removal of timepoints when sub-TR movement occurred (according to measures of error or movement, respectively). However, I also hypothesise improvements in sub-TR movement

artefact will reduce data quality, due to a lack of sensitivity to detect task-related activation associated with task-correlated motion. Instead, I hypothesise data quality will be highest for the Motion Fingerprint condition, as the method designed to cope with artefact in the presence of task-correlated motion, and as the method which can account for the widest range of movement artefacts with the fewest regressors in the GLM (relative to robust WLS and scan nulling).

8. METHODS

8.1 Sample

The whole sample (N=43) was included as in Experiment 1.

8.2 Language fMRI

The effect of artefact repair was investigated for two tasks in the Panda Games battery, which differed in terms of artefact severity; PD (high artefact) and LN (low artefact). The exclusion of RN and PN from this analysis served to reduce the number of statistical comparisons made. Further, because artefact severity for PN and RN was statistically equivalent to LN, results from this analysis can be generalised to these conditions.

8.3 Post-processing artefact repair methods

8.3.1 Standard GLM

Preprocessed images were entered into the GLM and effects of interest were modelled along with six realignment parameters, as described in Chapter 3. From here on this method is referred to as the standard GLM.

8.3.2 Expanded movement regressors and scan nulling

Preprocessed images were entered into the GLM along with 24 realignment parameters using a volterra expansion and a variable number of scan nulling regressors, dependent upon the number of volumes with movement >3mm (1 voxel). All other settings remained the same as the standard GLM pipeline.

8.3.3 Robust WLS

Robust WLS was performed using the spm8 toolbox (http://www.icn.ucl.ac.uk/motorcontrol/imaging/robustWLS.html). Unsmoothed data were entered into the standard GLM, to provide more independent data points when estimating the variance of residuals (Diedrichsen & Shadmehr, 2005). During calculation of the GLM a restricted maximum likelihood (ReML) approach was used to account for non-sphericity, producing unbiased estimates of residual variance. These estimates were then inverted, and used to de-weight noisy time points during WLS estimation of the beta weights. Beta images were smoothed (with a FWHM of 8) before producing contrasts of interest.

8.3.4 Motion fingerprint

The motion fingerprint method was run using the motion finger print toolbox (http://www.medizin.uni-tuebingen.de/kinder/en/research/neuroimaging/software/). Three motion fingerprint regressors were calculated for each participant (as described previously). Time-shifted versions of each regressor were also calculated, as per the Volterra expansion described previously, and with the aim of accounting for spin-excitation history effects (Friston et al., 1996). Inclusion of these additional shifted regressors improves performance of this method, by increasing the amount of variance explained by the GLM and thus reducing the residuals (Wilke, 2012). As such, in total six regressors were added as covariates of no interest into the first-level GLM for each participant.

8.4 Measures of scan quality

8.4.1 Subjective measure: Artefact ratings

Subjective artefact ratings from experiment 1 were used as a subjective measure of data quality in experiment 2, where ratings of 3 = poor quality, 2 = acceptable and 1 = good quality.

8.4.2 Objective measure: F_{adj}

Zsoter and colleagues provide an automated method for the objective classification of scan quality, based on the assumption that successful scans activate target regions (i.e. the language network) more than the rest of the brain (the control region) (Zsoter, Staudt, & Wilke, 2012). This algorithm produces a classification factor (F_c), which reflects the ratio of activated voxels inside and outside target regions, at a given statistical threshold, calculated as:

F = (sensitivity) / (false positive rate)

Where:

Sensitivity = (true positives / total target ROI voxels)

False positive rate = (false positives / total control ROIs voxels)

This method rewards sensitivity and penalises images based on the false positive rate. However, this can result in extremely high F_c values when both the false positive rate and sensitivity are low (the image is rewarded for having a low false positive rate but is not penalised for low sensitivity). In this situation F_c does not hold face validity compared to the gold standard for quality assessment (visual inspection): Despite the low false positive rate, an image with such poor sensitivity (a very high false negative rate) would typically be rated as poor quality. For group level analyses false negatives are a large concern. As such, I calculated an adjusted Fc (F_{adj}) according to:

F_{adj} = (sensitivity * specificity)

Where:

Sensitivity = true positives / true positives + false negatives

Specificity = true negatives / true negatives + false positives

The adjusted F (F_{adj}) rewards sensitivity and specificity with equal weighting, penalising images with either low sensitivity or low specificity. For the

purposes of this analysis F_{adj} was multiplied by 100 to avoid floor effects (sensitivity tended to be very low due to the use of very large target ROIs; especially for PD).

8.5 Regions of interest

8.5.1 Target ROIs

8.5.1.1 Listen and Name

The target region for LN (compared to its sensory baseline task, AG) was the whole auditory comprehension network, as defined by a subgroup activation map (LN>rest, $p<0.001$, k>10) (Figure 28). This included bilateral STG, left temporal pole, left Pars Triangularis (BA 45), left hippocampus and parahippocampal gyrus, right hippocampus, left thalamus, globus pallidus and SMA (significant on the group map at $p=0.05$, FDR corrected), and bilateral precentral gyri (motor cortex), left Pars Orbitalis (BA 47), right temporal pole, left cerebellum, bilateral rolandic operculum (significant on the group map at $p<0.001$ uncorrected, $k>10$). Using this broad target ROI, activation within any part of the auditory comprehension network was classed as true positive. This was seen as beneficial as the relative extent of activation associated with semantic and syntactic processing was hypothesised to vary between participants, within this broader network (for example due to age-related differences or level of engagement with the baseline task). The subgroup used to create this ROI included only those participants with clean (high quality) data, to reduce the influence of type I and type II errors on creation of the target ROI. Participants with any sub-TR movement artefact (N=21), movement > 1 voxel (3mm) between scans (N=2) or with task performance <60% accuracy (N=0) were excluded. Participants in the subgroup (N=20) were representative of the whole sample, and did not differ from excluded participants in terms of ability (IQ, core language score or in-scanner task performance). They were, however, significantly older (mean age = 11.59 years, SD = 3) and less strongly right hand dominant (mean Edinburgh handed score = 34.53, SD = 74.13) compared with excluded participants (mean age =

8.98 years, SD = 3.48; mean Edinburgh handedness score = 65.83, SD = 36.44). Encouragingly, the increased incidence of atypical handedness in this subgroup suggests the target ROI will account for some variation associated with atypical language dominance. Further, the group map showed biologically plausible language regions (Price, 2012) which were comparable to regions found previously in children using a similar task (Berl et al., 2013).

8.5.1.2 Picture Describing

To create the target region for PD (compared to its sensory baseline task, CNs) I first created an optimised group map (as for LN). Participants with high quality data (N=15, 8 females, 5-16 years) were representative of the whole sample, and did not differ from excluded participants in terms of age, ability or handedness. This group map showed significant clusters of activation (p<0.05, FDR) in bilateral inferior and middle occipital gyri, bilateral inferior temporal gyrus and fusiform gyrus, and bilateral primary motor cortex. At a lower threshold (p<0.001 uncorrected, k>10) activation was also seen in the right parahippocampal gyrus and right cerebellum. Previous studies of overt sentence generation in adults report activation beyond this network, including classic frontal and temporal language regions (reviewed in Chapter 5, section 3.2.2). Due to the small number of subjects with high quality data for this task and potentially large inter-subject variability, this comparison may have lacked statistical power to detect higher-order cortex in the sentence generation network. As such, for PD I used a hypothesis-driven anatomical ROI as the target region. This anatomical ROI was formed of all regions identified in studies of overt sentence generation in adults (Chapter 5, section 3.2.2). Both left and right hemisphere homologues of each region were included in the ROI, to account for atypical language lateralisation (Springer 1999) and the more bilateral distribution of language in children compared to adults (Berl et al., 2014; Szaflarski, Holland, et al., 2006). The target ROI for PD (Figure 28) therefore included: IFG (BA 44,45,47), the occipito-temporal region including fusiform gyrus (BA 37), rolandic operculum, STG/STS/MTG (BA 21 and 22), caudate nucleus and putamen, primary and pre-motor cortices (BA 4 and 6), somatosensory cortex (BA 1, 2, 3), pre-SMA and planum temporale. These

regions were selected from the automatic anatomical labels (Tzourio-Mazoyer et al. 2002) available in the Wake Forest University Pick Atlas (Maldjian et al. 2003; Maldjian et al. 2004), and were combined using MarsBar (Brett, Anton, Valabregue, & Poline, 2002).

8.5.2 Control ROIs

The control ROI for each task was based on all voxels within the whole brain template provided by spm8, minus the target ROI.

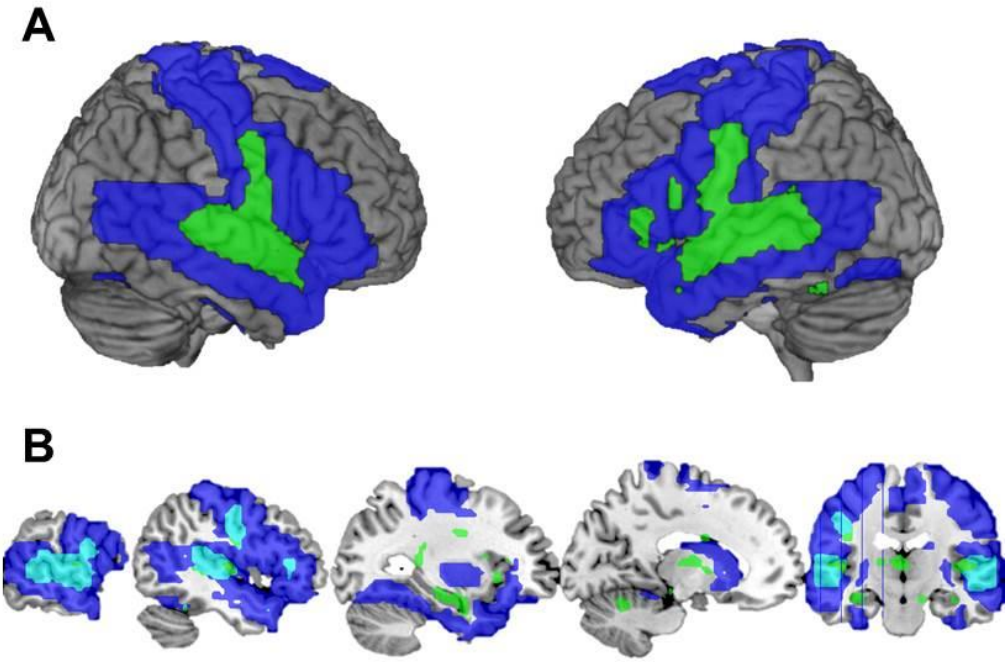


Figure 28: Target ROIs for LN (green) and PD (blue) on a rendered brain (A), and on sagittal slices (B) demonstrating overlap of ROIs (cyan).

9. RESULTS

9.1 Effect of repair on artefact ratings

For both tasks, severe artefact was least common and artefact-free data most common following repair using robust WLS (Table 13). However, none of these differences were statistically significant when compared using Friedman's ANOVA.

	Listen and Name			Picture Describing		
Method	**None**	**Mild**	**Severe**	**None**	**Mild**	**Severe**
Standard GLM	51%	35%	14%	44%	28%	28%
Scan nulling	50%	31%	19%	41%	24%	36%
Robust WLS	61%	37%	2%	61%	30%	9%
Motion fingerprint	49%	37%	14%	37%	33%	30%

Table 13: Subjective artefact ratings (none, mild, severe) for two language tasks, following artefact repair using four methods: Standard general linear model (GLM), expanded movement regressors and scan nulling (scan nulling), robust weighted least squares (robust WLS) and the motion fingerprint.

9.2 Effect of repair on scan quality (F_{adj})

9.2.1 Validation of F_{adj}

In order to assess the face validity of F_{adj}, seven images (individual results for LN>rest) were ranked based on F_{adj} and according to visual quality judgements by a reviewer blind to F_{adj}. To ensure a cross-section of images were rated, images were selected to represent each of the following percentiles based on F_{adj}: 5^{th}, 10^{th}, 25^{th}, 50^{th}, 75^{th}, 90^{th} and 95^{th}. The rater was instructed to list images from 1 to 7 in order of quality, where 1is the highest quality image, and according to set criteria (see Appendix 9.). The contrast LN>rest was used for this comparison, to provide more reliable visual inspection results; the contrast LN>AG is experimental and less research exists on which to base visual quality judgements. Figure 29 shows good correspondence between objective (A) and subjective (B)

measures of scan quality at the individual level, suggesting F_{adj} has face validity. The only discrepancy occurred for rating of poor quality images (F and G), where F_{adj} penalised false positives over false negatives while visual inspection did the reverse.

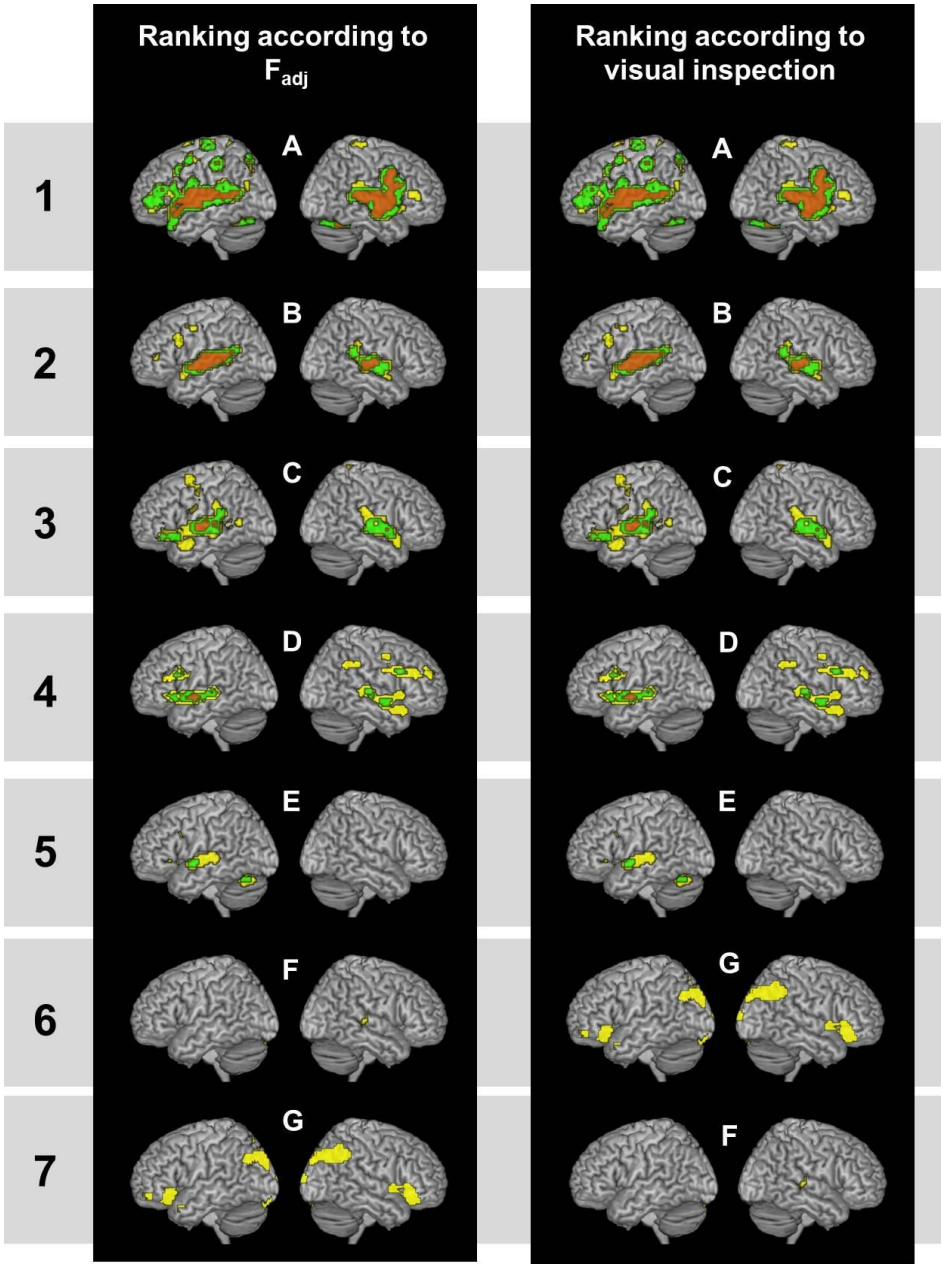


Figure 29: Validation of an objective measure of scan quality (F adjusted) against the current gold standard in scan quality assessment; visual inspection. Images were ranked according to quality as assessed by the objective measure (F adjusted) and by subjective (blind) visual inspection (right), showing strong agreement between methods. Spm *t* results are displayed at three statistical thresholds; $p<0.05$ FWE corrected (red), $p<0.001$ and $k>20$ (green), and $p<0.01$ and $k>20$ (yellow).

9.2.2 Effect of repair

Due to the large SD and influence of outliers on $F_{adj,}$, statistical analyses were performed using non-parametric tests (Friedman's ANOVA, Mann-Whiteney U or Wilcoxon Signed-Ranks test), and the trimmed mean (which excludes the upper and lower 5% of data, as a means to remove the influence of outliers) and 95% confidence intervals are presented, to aid interpretation (Table 14). For LN, there was no significant difference in F_{adj} between participants with and without artefact for any repair method. For PD and the motion fingerprint repair method, there was a trend for significantly increased F_{adj} in patients with artefact relative to those without artefact (p=0.07). Based on these findings, method comparisons were performed in all participants, regardless of artefact severity.

	F (adjusted)					
	Whole sample (N=43)				*Artefact groups*	
	Mean	**SD**	**Trim.**	**95% CI**	**-**	**+/++**
LN>AG					***N=22***	***N=21***
Standard GLM	9.45	2.22	7.05	4.97 - 13.93	7.44	7.62
Scan nulling	7.40	1.54	6.06	4.29 - 10.51	4.85	7.33
Robust WLS	0.71	0.53	0.16	-0.36 - 1.77	0.15	0.17
Motion fingerprint	10.27	2.28	7.97	5.67 - 14.86	7.07	9.26
PD>CNs					***N=19***	***N=24***
Standard GLM	3.07	0.56	2.65	1.93 - 4.20	2.51	2.78
Scan nulling	2.04	0.43	1.72	1.18 - 2.90	2.06	1.45
Robust WLS	0.10	0.03	0.06	0.03 - 0.16	0.04	0.08
Motion fingerprint	2.73	0.56	2.30	1.60 - 3.86	1.20	3.18

Table 14: Descriptive statistics for F_{adj} are presented for an auditory comprehension task (LN>AG) and sentence generation task (PD>CNs), across four repair methods: standard general linear model (Standard GLM), expanded movement regressors and scan nulling (scan nulling), robust weighted least squares (Robust WLS) and motion fingerprint. Trimmed means (Trim.) are presented for the whole sample and for participants without (-) and with minor (+) or severe (++) artefacts, as are 95% confidence intervals (CI), due to the large standard deviation (SD) and influence of outliers.

There was a significant difference in F_{adj} between methods for both the low artefact (LN) and high artefact (PD) tasks ($\chi^2(3)=72.58$, $p<0.001$, and $\chi^2(3)=71.68$, $p<0.001$, respectively). Post-hoc comparisons with a Bonferroni correction for multiple comparisons ($\alpha=0.008$) showed significantly lower F_{adj} for robust WLS relative to every other method, for both tasks (all p values <0.001). Further post-hoc analysis showed that for both tasks, the true positive rate was significantly lower for robust WLS compared to all other methods ($p<0.001$), suggesting low F_{adj} results from reduced sensitivity of robust WLS versus all other methods. For LN, there was also a trend such that scan nulling produced a lower F_{adj} statistic relative to the standard GLM ($p=0.06$).

10. DISCUSSION

10.1 Prevalence of speech-related movement artefact in the Panda Games battery

Results from Experiment 1 show speech-related movement artefacts occur in approximately 40% of cases for the Panda Games task battery. This high prevalence rate is due to the use of an overt speech response mode, combined with ascending image slice acquisition. When slices are acquired in the axial plane and in ascending acquisition mode, head movements during speech occur within the slice acquisition plane, meaning that movements larger than one voxel can result in whole slices being displaced. Displacement of a slice can result in axial signal dropout (as shown in Figure 26). At the group level, this could cause an inflation of false negatives (type II error) should regions of signal drop occur consistently across individuals. Encouragingly, this form of artefact varies between participants. As such, drop out at the group level should be minimal for participants with minor artefact (drop out in a single slice of part of a slice). Further, the group mask can easily be checked for any regions of drop out. Inclusion of participants with severe artefact (drop out occurring across multiple slices) could result in overlapping regional drop out across participants. As such, these participants will be excluded from analyses in future chapters. The presence of severe artefact was much less common, and occurred in 9%, 28%, 14% and 15% of participants for PD, PD, LN and RN, respectively. For

word level tasks (PN, LN and RN), these failure rates are lower than would be expected as a consequence of in-scanner movement; ~16% in previous studies of healthy children (Yerys et al., 2009). The failure rate for PD is almost twice as high. However this was expected for an experimental, whole-sentence generation task. Further analyses in Experiment 1 suggest this type of movement artefact is more common in young children, <8 years of age. There were no differences in artefact severity between males and females, suggesting both are vulnerable to this type of artefact. Whether or not the use of mock scanning has an effect of reducing in-scanner movement remains to be seen in Chapter 5.

10.2 The optimal method for repair of sub-TR movement artefact

On average robust WLS showed the greatest improvement in sub-TR movement artefacts, although this was not a statistically significant difference. On average the prevalence of minor or severe artefacts was reduced by 10% for LN and 17% for PD, when compared to the standard GLM. This improvement reflected a decrease in severe artefact relative to the standard GLM, from 14% to 2% in LN and from 28% to 9 % in PD. Robust WLS de-weights noisy time points during calculation of the beta weights. The findings from Experiment 1 suggest this is the most effective strategy for removing sub-TR movement artefact on an individual basis. Scan nulling works on a similar principle; nulling the influence of certain time points during calculation of the beta weights. However, with this method noisy time points are defined by scan-to-scan movement over a specified threshold (3mm/TR in this study), rather than the amount of error (unexplained variance). Scan nulling did not improve sub-TR movement artefacts, suggesting scan-to-scan movement (at least at the threshold >3mm) does not adequately capture sub-TR movement. It may be that movements <3mm can also cause sub-TR movement artefacts, or that movements >3mm don't always cause this type of artefact. Other aspects of in-scanner movement may be of more importance to axial signal dropout, rather than size of movement alone, such as the direction or timing of movement. One might hypothesise that rotations are specifically detrimental, due to the interaction of movement with B_0 (as discussed in section 3.1.4). The Inclusion of the standard realignment parameters, expanded realignment parameters and even the inclusion of

parameters which account for B_0*movement interactions (the motion fingerprint) fail to repair sub-TR movement artefact. These findings suggest that whole volumes must be removed or de-weighted to remove this artefact, until methods are developed to identify and repair outlying time points on a voxelwise basis. Initial findings from the comparison of methods presented here, suggest identifying and repairing noisy time points based on their error variance may be an effective strategy. Crucially, whether removal of this type of artefact translates into higher quality scans is an important consideration, and is discussed below.

10.3 Development and validation of an objective measure of scan quality

In order to compare artefact repair methods and to quantify the effect of repair, I developed an objective and automatised measure of scan quality (F_{adj}) adapted from the classification factor (F_c) introduced by Zsoter and colleagues. The original classification factor was calculated according to both sensitivity and specificity, but rewarded specificity over sensitivity. Although this provides stringent control over false positives, it also results in poor face validity when compared with the current gold standard in scan quality assessment (visual inspection by an expert reviewer). Most notably, a scan with very low activation might obtain a high F_c statistic due to the specificity of findings to the target network, without taking into account the lack of activation in expected regions. To address this issue, I adjusted F such that sensitivity and specificity were given equal weighting. The F_{adj} statistic therefore penalises scans when either specificity or sensitivity are low. When compared with visual ratings, F_{adj} shows good face validity as a measure of scan quality at the individual level (Figure 29).

10.4 The optimal method for improving scan quality

In Experiment 2, I used F_{adj} to assess the impact of artefact repair on individual data quality; crucial in the pre-surgical setting. Findings showed that while robust WLS provides the best repair of sub-TR movement artefact, this comes at a cost. Scan quality was strikingly poor for robust WLS compared to all other methods, achieving an F_{adj} of only 0.16 for LN and 0.06 for PD, compared to 7.05 and 2.65 for the

standard GLM. Post-hoc comparisons of the true positive rate showed robust WLS was associated with significantly lower sensitivity compared to all other repair methods. Thus, de-weighting noisy volumes reduces sensitivity to detect task-related activation. This is most likely due to the strong correlation between sub-TR movement and task performance, such that volumes selected for de-weighting during robust WLS are therefore the volumes, which contain task-related changes in the BOLD response. There was also a trend for poorer data quality as a result of repair by scan nulling, which works on a similar principle to robust WLS. Both findings emphasise the need for voxelwise repair of sub-TR movement artefact, which can preserve task-related BOLD changes occurring in the same volume.

Both the standard GLM and motion fingerprint methods produced statistically equivalent data quality. Unlike the standard GLM, the motion fingerprint accounts for spin excitation history effects and movement*B_0 interactions, as well as scan-to-scan movement. Crucially, it accounts for these using minimal (six) regressors in the GLM, increasing its sensitivity relative to other methods, which use additional regressors to account for these effects (i.e. realignment parameters and scan nulling regressors). Encouragingly, the similarity of data quality between the standard GLM and motion fingerprint suggests the influence of more complex movement artefacts (not accounted for by the standard realignment parameters) is small in this cohort. When comparing trimmed means and 95% confidence intervals however, the motion fingerprint method did produce higher quality data on average, such that 95% of participants obtained an F_{adj} value between 5.67 and 14.86 for LN, and 1.60 and 3.86 for PD (although this was not a statistically significant difference). The motion finger print method therefore accounted for some additional variance in the GLM relative to the standard method on an individual basis, albeit a small amount.

11. CONCLUSIONS

The results of Experiments 1 and 2 suggest sub-TR movement artefacts are best repaired by de-weighting high-error time points during calculation of beta weights (for example using robust WLS). However, this also reduces task related activation in the context of task-correlated movement, rendering this method invalid for use in future

chapters. In contrast, the motion fingerprint method improves data quality (on average) by accounting for more complex effects of movement relative to the standard six realignment parameters. It does not remove sub-TR movement artefact however. Although this type of artefact causes false negatives at the individual level, this will not translate to false negatives at the group level unless signal dropout occurs consistently across individuals. Participants with severe artefacts are more likely to show overlapping regions of signal dropout, and so will be excluded from analyses in future chapters. Participants with minor artefacts however will not create false negatives at the group level. As such, these participants will be included in future chapters, should they adhere to additional data quality exclusion criteria as per Chapter 5.

PART 3: VALIDATION

Chapter 5: Validation of the Panda Games in healthy children and adolescents aged 5-16 years.

1. ABSTRACT

Before the new Panda Games fMRI task battery can be used in the presurgical setting, the tasks must be validated in healthy children of the same age range as surgical candidates. In this chapter I aimed to; define the typical networks supporting performance of the Panda Games, assess scan failure rates in children, evaluate the effectiveness of pre-scan preparation with a mock scanning procedure, and investigate how well age-appropriate task design controls for performance accuracy across age. Forty three healthy children (5-16 years) performed the Panda Games task battery during fMRI scanning. Age, sex and anxiety-matched children who did (N=13) and did not (N=13) undergo mock scanning were compared. All tasks produced activation in expected frontal and temporal regions. Similar to adults, amodal semantic processing was supported by left Pars Triangularis, temporal pole and ventral occipito-temporal cortex, as well as bilateral anterior fusiform gyrus. I hypothesise this ventral network may be particularly crucial for postoperative language outcome in children with epilepsy, but this requires future investigation. Scan failure rates ranged from 16% - 41% (similar to previous studies) and were highest for the sentence generation task. Movement was the largest cause of scan failure, and was associated with boredom and discomfort during scanning. Children who underwent mock scanning performed more accurately inside the scanner. I conclude that the Panda Games task battery successfully localises language networks in children as young as 5 years of age, and that pre-scan preparation can be helpful for obtaining good quality fMRI data in young children.

2. AIMS

In order to validate the Panda Games task battery for use in healthy children, I aim to:

1. Summarise scan failure rates associated with the Panda Games, and identify the principal causes of scan failure to be targeted in future work/in the clinical setting.
2. Evaluate the effectiveness of mock scanning on measures of scan quality (in-scanner movement, task performance and scan failure rates).
3. Define typical networks supporting performance of each task in children and adolescents, including localisation of core linguistic functions (semantic and syntactic processing) and non-linguistic (supporting) functions (such as sensory-motor processing and cognitive control).

To provide new information regarding the organisation of the language system (relevant to pre-surgical neuroimaging), I also aim to identify regions which are consistently activated across the Panda Games, associated with amodal semantic processing.

3. INTRODUCTION

Before the Panda Games can be used in the clinical setting, it is important to validate their use in typically developing children of the same age as surgical candidates (5-16 years). Most importantly, to assess 1) failure rates associated with each task and to identify participant characteristics predictive of scan failure, 2) the clinical usefulness of preparing children for their scan using mock scanning, and 3) the typical networks engaged by each task, in order to provide a typical baseline to which children with epilepsy can be compared, and from which predictions can be made.

3.1 Obtaining successful scans in paediatric populations

For paediatric populations, the success rate of MR imaging increases with age (Byars et al., 2002; Yerys et al., 2009) as key executive functions, such as the ability to inhibit movement and sustain attention, develop. As such, very young children (< 7 years) show very high failure rates, of up to 57% (Byars et al., 2002). Failure rates are also particularly high when children are asked to perform a whole battery of tasks

during scanning: 31% for healthy children and 41-50% for children with neurological conditions such as epilepsy (Yerys et al., 2009). However, research suggests age-related differences in scan failure rates can be removed when appropriate pre-scan preparation protocols are put in place (de Bie et al., 2010). Thus, many studies have begun to adapt their MR protocol to improve scan success rates in paediatric populations (Berl et al., 2014; de Bie et al., 2010; Hallowell et al., 2008; Raschle et al., 2009; Schlund et al., 2011; Thomason, 2009). Based on these previous studies, I developed a pre-scan preparation protocol (Chapter 3), which also included mock scanning for young or anxious participants. This protocol was designed to target principle causes of scan failure in paediatric populations; poor task performance (which accounts for up to 60% of scan failures; (Croft et al., 2013), anxiety (up to 57%; Bryars et al., 2002), and in-scanner movement (~16% of scan failures; Yerys et al., 2009; Croft et al., 2013). Although the effect of pre-scan preparation cannot be fully quantified in this study (all participants took part in the protocol) participation in mock scanning was variable (dependent on the participant's age and needs). As such, I aim to compare scan failure rates and the principle causes of scan failure (in-scanner movement, anxiety and task performance) between those who did and did not undergo mock scanning, when matched for age and ability. I hypothesise the mock scanning group will show reduced scan failure rates compared to participants in the no-mock group, reflected in reduced in-scanner movement, lower anxiety and better task performance.

3.2 Language networks supporting the Panda Games

The Panda Games were designed to engage four language networks; word retrieval (Picture Naming; PN), sentence generation (Picture Describing; PD), auditory comprehension (Listen and Name; LN) and reading comprehension (Read and Name; RN). By probing both expressive (PN and PD) and receptive (LN and RN) networks, at both the word (PN) and sentence level (PD, LN and RN), this battery aims to provide comprehensive mapping of the entire language network, such that pre-surgical neuroimaging can be adapted to each individual patient's needs (i.e. age and ability, location of surgical target, planned surgical procedure, and the specific surgical question). In the following sections I provide hypotheses for networks supporting each task in the Panda Games battery, based on the functional

neuroanatomical model of language proposed by Price (2012) (Figure 30 and Figure 31). This model is based on the most consistent structure-function mappings evident from the last 20 years of fMRI and positron emission topography (PET) investigations of language, and builds on an earlier model of lexical processing by Petersen (Petersen, Fox, Posner, Mintun, & Raichle, 1988; Petersen et al., 1989). Because this model provides a comprehensive summary of consistent findings from a vast literature, I use this model as the basis for my interpretation of group-level findings. It is therefore referenced throughout this chapter. Specific studies are referenced where they provide additional information. I also conducted targeted literature reviews to supplement my hypotheses relating specifically to the Panda Games and to the paediatric population, as the functional neuroanatomical model does not include predictions for sentence-level (syntactic) processing, and is primarily based on studies of the mature language system.

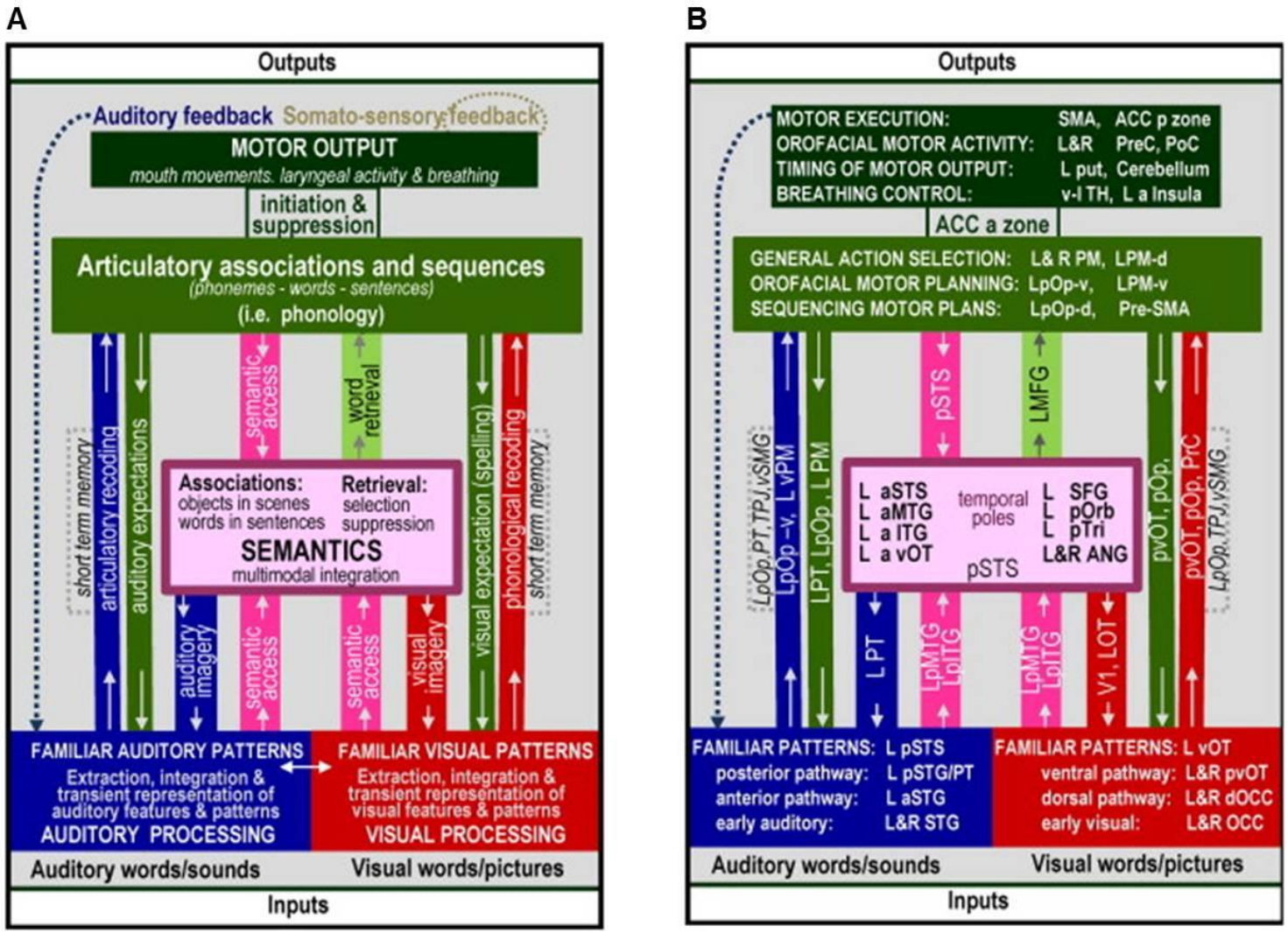


Figure 30: Processing stages (A) and the neural substrates (B) of the Functional Neuroanatomical Model of language proposed by Price (2012). This model is based on consistent structure-function mappings from the last 20 years of neuroimaging investigations of language using fMRI and PET. For an explanation of abbreviations used in the model see Figure 31 overleaf. This figure is taken from Price 2012.

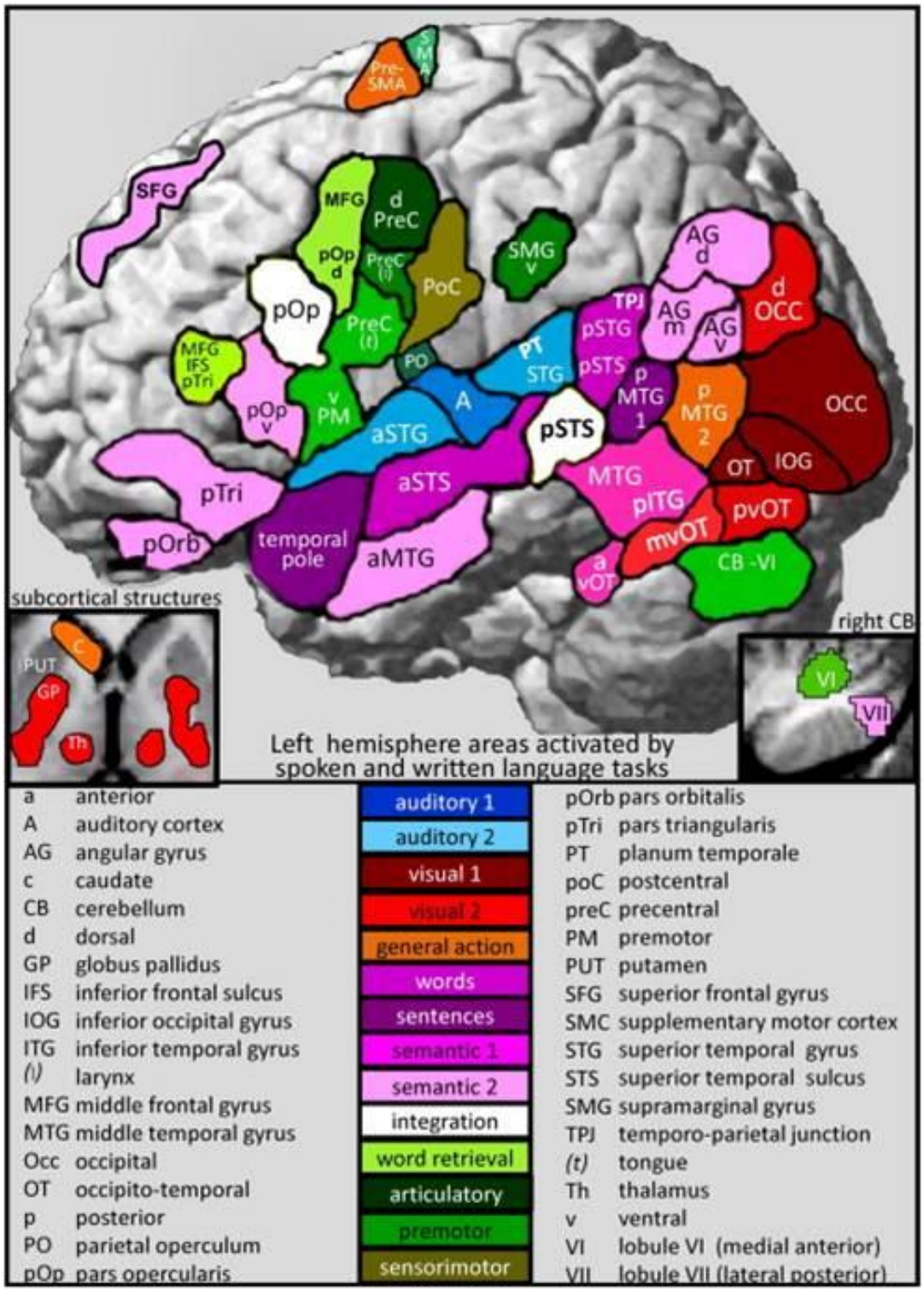


Figure 31: The neural substrates of the Functional Neuroanatomical Model of language, proposed by Price (2012). A cartoon representation of the most consistent structure-function mappings found using fMRI and PET in adults. Processing stages are colour coded, and anatomical regions are labelled. This figure is taken from Price 2012.

3.2.1 Picture naming: Semantically-based word retrieval

According to the functional neuroanatomical model of language (Price 2012) word generation (in response to a visual stimulus) first involves visual processing in occipital and ventral occipito-temporal cortex. This is followed by gaining semantic access (left posterior MTG/ITG) through which the semantic associations of the visual stimulus can be obtained (left temporal lobe: anterior STS/MTG/ITG and ventral occipital-temporal cortex). Words can then be retrieved in order to convey the formed concept, which is supported by frontal and parietal regions (left SFG, BA 45/47, bilateral AG). The relevant articulatory associations are then retrieved, including action selection (bilateral premotor cortex, left dorsal premotor cortex), orofacial motor planning (left ventral BA 44 and premotor cortex), and sequencing of motor plans (left dorsal BA 44 and pre-SMA). The response is then initiated or suppressed (ACC). If a response is initiated, this involves motor execution (SMA and ACC), orofacial motor activity (bilateral pre- and postcentral gyri), timing control over motor output (left putamen and cerebellum) and breathing (ventral thalamus and left anterior insula). The expected and actual speech output are associated with further auditory processing (STG/STS and PT), with articulatory recoding (left BA 44 and ventral premotor cortex) – supported by short term memory (BA 44, PT, SMG) – adjusting the speech stream as necessary. This process works very much in a loop system, with speech adjustments causing further auditory feedback processing and adjustments etc. This model involves a wide range of brain regions in single word generation.

Because the baseline condition (Colour Naming with words) is designed to involve visual processing, word retrieval, visual imagery and the same articulatory processes as Picture Naming, a contrast of Picture Naming > Colour naming (with words) should localise regions supporting semantic access and association. In adults these include: left anterior STS/MTG/ITG, posterior MTG/ITG and ventral occipital-temporal cortex (vOT). Studies of word retrieval in children (reviewed in the introduction of this thesis) suggest activation of these regions will be weaker (Szaflarski, Holland, et al., 2006; Szaflarski, Schmithorst, et al., 2006;

Vannest et al., 2010) and more bilateral (Everts et al., 2009; Holland et al., 2001; Karunanayaka et al., 2010; Szaflarski, Holland, et al., 2006; Szaflarski, Schmithorst, et al., 2006) in younger children.

3.2.2 Picture describing: Sentence generation

Although the functional neuroanatomical model of language (Price 2012) proposes a model for word generation, few studies have been conducted to investigate overt *sentence* generation. This may be due to some of the methodological difficulties associated with imaging overt speech (Chapter 1). Despite this, there are a small number of studies already published in this area, and as scanning and post-processing methods develop to deal with movement issues (Chapter 4), this area of research is likely to grow. Earlier studies used PET, while more recently sentence generation is being investigated using fMRI.

In 1999 Levelt proposed a cognitive model of sentence generation - 'the blueprint of the speaker' (Levelt, 1999) (Figure 32) - which provides a useful framework within which to review previous studies of overt sentence generation. The following paragraphs provide an overview of studies investigating each processing stage described by this model. I then summarise my hypotheses for expected activation during the Picture Describing task in children and adolescents.

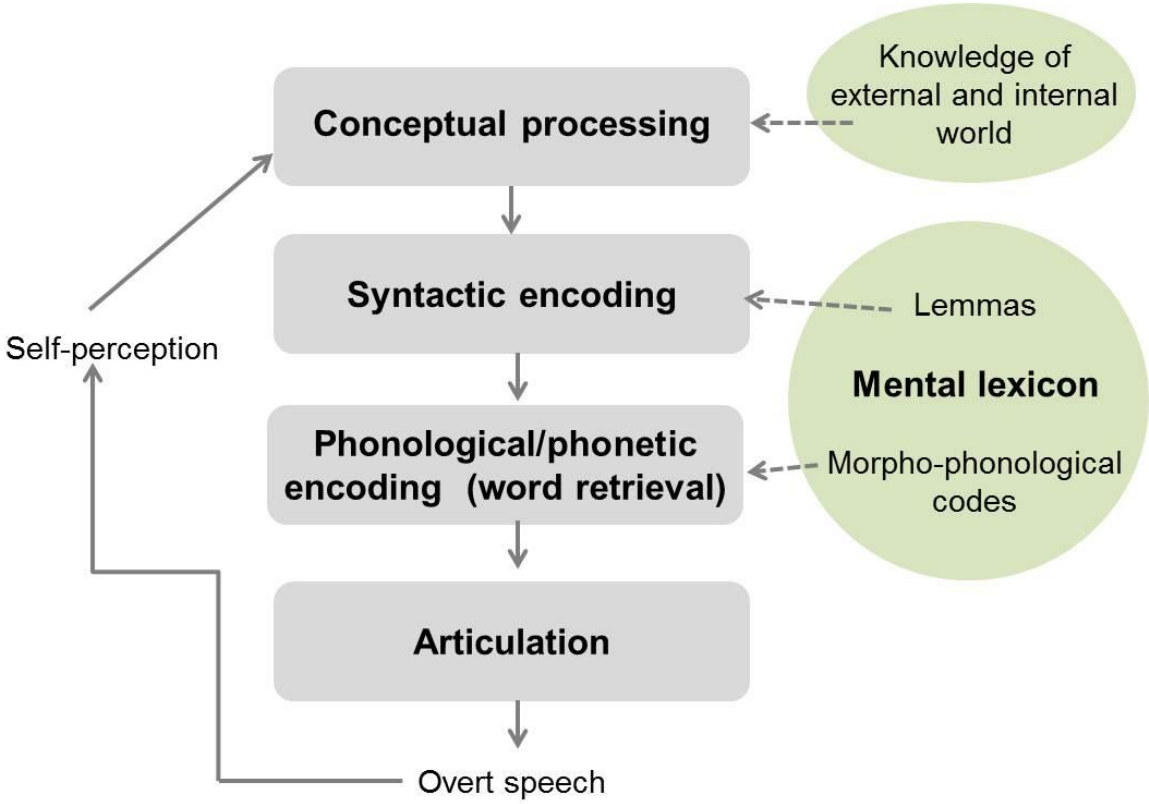


Figure 32: Simplified reconstruction of the 'blueprint of the speaker'; a model of sentence production proposed by Levelt (1999). Grey boxes represent principal processing stages in the model.

3.2.2.1 Conceptual processing

In Levelt's model, the conceptual processing stage involves the generation of a message, which is designed to have a specific effect on the listener, in order to reach a goal. This stage relies on knowledge of the external and internal world. However, in this review, I consider the conceptual processing stage in Levelt's model to reflect semantic processing, and here review studies which have tried to localise these processes. Such studies typically compared overt sentence generation to an overt baseline task with low semantic and syntactic processing load (e.g. sentence repetition, production of overlearned nursery rhymes or counting). The number of published studies is limited (five), as are their sample sizes (N=8-21). Further, no studies could be found with children or adolescents. Despite this, there are some consistencies across studies, which I

can use to guide my hypotheses regarding the expected activation pattern for the Picture Describing task. Most notably, vOT was identified across all studies. In the context of responding to a visual picture, this region is specifically implicated in semantic associations in the functional neuroanatomical model of word production.

The response of the frontal lobe was more variable across studies. Some studies reported variable activation of BA 44, 45 and 47 (Awad, Warren, Scott, Turkheimer, & Wise, 2007; Dhanjal, Handunnetthi, Patel, & Wise, 2008; Tremblay & Small, 2011), while others reported no activation in these regions (Blank et al., 2002; Muller et al., 1997). Within Broca's area, BA 45 was most consistently activated. Two studies also reported activation superior and anterior to regions typically associated with single word generation, such as the dorsolateral prefrontal cortex (DLPFC) (Muller et al., 1997) and orbitofrontal cortex (Awad et al., 2007). Together these regions have been implicated in reasoning and episodic memory, with the DLPFC evaluating externally generated information and the orbitofrontal region evaluating internally generated information (Christoff et al., 2001). Findings from studies of sentence generation reviewed here provide support for this account: DLPFC activation is seen for externally generated sentences (based on an example sentence, and using a target word) (Muller et al., 1997), while activation of orbitofrontal cortex is seen for internally generated sentences (based on episodic memory) (Awad et al., 2007). These more extensive frontal regions therefore appear to reflect metalinguistic processes during sentence generation, and fit nicely with Levelt's proposal that the conceptual processing stage relies on internal and external representations of the world (Figure 32).

Other regions which were inconsistently active across studies included: ventro-lateral anterior temporal lobe (Blank 2002; Dhanjal 2008), pre-SMA (despite baseline conditions with articulation) (Blank et al 2002; Awad et al., 2007; Tremblay and Small 2011) and the angular gyrus (Muller 1997; Blank 2002). Inconsistencies in findings may have arisen due to differences in methods (PET and fMRI), small sample sizes and task differences (for example, there was

often little description of the baseline condition and its associated activation pattern).

3.2.2.2 Syntactic encoding

During this stage the syntactic form of the relevant concepts (the 'lemmas') are retrieved from the mental lexicon and ordered to give the sentence its surface structure. Two studies investigated syntactic processing during sentence generation (Haller, Radue, Erb, Grodd, & Kircher, 2005; Indefrey et al., 2001). Indefrey and colleagues compared three tasks, which differed in syntactic demands; describing a picture with a) sentence, b) noun phrase, c) a string of single (disjointed) words. A graded response was found in the left rolandic operculum (just posterior to Broca's area), with strongest activation for sentences (highest syntactic load) and weakest activation for disjointed words (lowest syntactic load). Haller and colleagues compared activation during sentence generation in response to three visually presented words, with activation during sentence repetition (from written stimuli). By providing participants with the subject, verb and object in the generation condition, the authors aimed to keep semantic processing demands minimal but syntactic demands high relative to the repetition condition. Comparing these two tasks produced activation in left IFG (BA 45>44) and left MTG. Increases in activation were also seen in the left superior parietal lobe (somatosensory association cortex, BA 7), SMA, right insula and right somatosensory cortex (BA 1 and 3). These regions have been associated with articulation and auditory feedback (Price 2012).

3.2.2.3 Phonological/phonetic encoding (lexical retrieval)

According to Levelt's model, the phonological form of a lemma is retrieved, and the associated articulatory code for each phoneme is obtained ready for articulation. The rate of words spoken (during a self-generated and self-paced narrative) was associated with activation in the STG/STS/MTG, which the authors interpret as reflecting lexical retrieval (Kircher, Brammer, Williams, & McGuire, 2000). The STG/STS is a common region of activation for propositional speech, counting and repeating nursery rhymes, and also for sentence generation and repetition (Blank et al., 2002; Tremblay & Small,

2011); all of which involve word retrieval and phonological processing. One study focused on the role of the basal ganglia in language production, and found the caudate nucleus to be specifically involved in sentence generation, and not repetition (Argyropoulos, Tremblay, & Small, 2013). The authors suggest the caudate is involved in semantic aspects of response selection, or cognitive control (i.e. inhibition of inappropriate responses and release of appropriate responses from inhibition).

3.2.2.4 Articulation

Several studies performed contrasts which allowed isolation of articulatory networks (Argyropoulos et al., 2013; Blank et al., 2002; Dhanjal et al., 2008; Gracco, Tremblay, & Pike, 2005; Tremblay & Small, 2011). Findings were more consistent for this network (harking back to the point made by Lieberman and colleagues regarding the bias for large effects in sensory-motor cortices in neuroimaging studies, as discussed in Chapter 3). Activated regions included; primary and pre-motor cortices (BA 4 and 6), somatosensory cortex, pre-SMA and putamen.

3.2.2.5 Prediction, feedback, error detection and repair

In his blueprint model, Levelt did not include a distinct 'module' relating to post-articulation processes such as somatosensory and auditory feedback, error detection and repair; all of which are now known to be of importance to speech production (Price, 2012; Price & MacSweeney, 2011). I have therefore included this section to review the few studies which have been able to address this using overt sentence generation tasks (Awad 2007; Dhanjal 2008; Tremblay & Small, 2011) or word retrieval tasks (Christoffels, Formisano, & Schiller, 2007; Tourville, Reilly, & Guenther, 2008). The most consistent finding across these studies is activation of the planum temporale in the STG.

3.2.2.6 Picture describing: Summary and hypotheses

Findings regarding semantic processing during overt sentence generation fall within the predictions of the functional neuroanatomical model, implicating left vOT (BA 37) and Pars Triangularis (BA 45). However, my review also identifies findings which are discrepant with the functional neuroanatomical

model. For example, word retrieval is associated with the middle frontal gyrus in the functional neuroanatomical model, but for overt sentence generation has been shown to involve the middle and superior temporal gyri. Because it is mainly word-based, the functional neuroanatomical model also has no specific 'syntax' module, while studies of sentence generation suggest syntactic processing may engage distinct regions. One study proposed the rolandic operculum as a potential substrate for syntactic processing (Indefrey et al., 2001), and further syntactic processing regions have been proposed based on studies of sentence comprehension (discussed later).

My hypotheses for expected activation during performance of PD are summarised in Figure 33. PD is a sentence generation task, which involves all stages of the sentence generation model proposed by Levelt. When compared to a resting baseline, PD should therefore engage a widespread network of brain regions including: IFG (principally BA 45), frontopolar regions, ventral aTL, STG/STS/MTG, rolandic operculum, IPL (principally the AG), primary and pre-motor cortices, pre-SMA, somatosensory cortices, the planum temporale, caudate nucleus and the putamen. The baseline condition (Colour Naming with Sentences) involves sentence generation with minimal semantic and syntactic load; the same sentence is repeated for each stimulus ("The colour of the square is…"), and only the final word (the colour name) must be retrieved. While both PD and CNs should activate brain regions associated with phonological processing, word retrieval, articulation and post-articulation processes, PD should produce a greater burden on semantic and syntactic processing regions. Based on studies of conceptual processing in sentence generation, I hypothesise activation in the left vOT and Pars Triangularis (BA 45) when PD is compared with CNs. Activation may also be seen in Pars Opercularis (BA 44) and Orbitalis (BA 47), orbitofrontal cortex, ventro-lateral anterior temporal lobes, pre-SMA and the angular gyrus. Based on studies of syntactic processing in sentence generation I also hypothesise activation in the rolandic operculum, left superior parietal lobe (BA 7) and left MTG. I could not find any studies which investigated sentence generation in children. As

such, my hypotheses are based only on this small selection of findings in adults, and my analysis of PD in children should be considered exploratory.

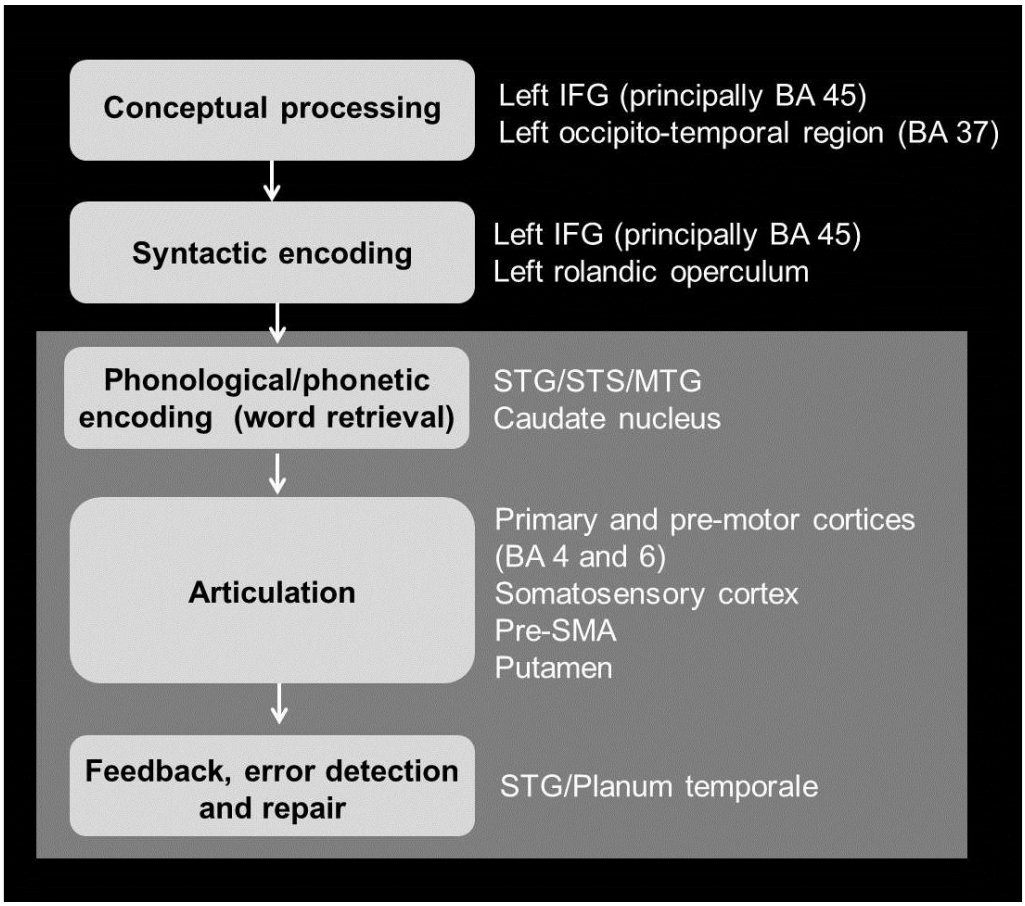


Figure 33: Hypotheses for expected activation during performance of the Picture Describing task. Expected regions of activation are listed (white text) for each processing stage in Levelt's (1999) 'blueprint of the speaker' model (light grey boxes). Processes and brain regions hypothesised for the Picture Describing task (PD) are within the black box, while processes and brain regions hypothesised for the Colour Naming with sentences baseline task (CNs) are within the dark grey box. By comparing PD and CNs, I aim to localise activation specific to PD (in the black box only) i.e. conceptual and syntactic processing stages. Expected regions of activation are hypothesised based on the most consistent findings from a small selection of studies investigating overt sentence generation in healthy adults using fMRI and PET.

Note the additional processing stage added to Levelt's model, to include feedback, error detection and repair processes into the expected activation pattern.

3.2.3 Listen and Name: Sentence-level semantic and syntactic processing from phonology

3.2.3.1 Auditory comprehension

The functional neuroanatomical model of language (Price, 2012) suggests auditory comprehension involves early auditory processing in bilateral STG, with familiar auditory patterns (words) specifically activating the left posterior STS. Semantic access is then gained via the left posterior MTG and ITG, or instead via the posterior STS based on the articulatory associations of heard speech (obtained via the left planum temporale, Pars Opercularis and premotor cortex). Semantic processing of the heard signal involves the pSTS (multimodal integration), anterior STS/MTG/ITG and vOT (semantic associations), and left Pars Triangularis (BA 45), Pars Orbitalis (BA 47), left SFG and bilateral AG (semantic retrieval). The appropriate lexical response is then selected (left MFG), prior to obtaining articulatory associations and sequences (left Pars Opercularis, bilateral premotor cortices and SMA), and motor output (bilateral pre- and post-central gyri, SMA, ACC, left putamen, anterior insula and thalamus). Any articulatory recoding that is required (due to an error for example) involves left Pars Opercularis and ventral premotor cortex, supported by short-term memory (left Pars Opercularis, planum temporale, temporo-parietal junction/SMG). Children engage a similar network during auditory comprehension, including: bilateral (left>right) STG, STS, MTG and fusiform gyrus, left AG, IFG/MFG and precuneus/posterior cingulate, bilateral SFG, hippocampi, precentral gyri and SMA (Ahmad et al., 2003; Berl et al., 2010; Lidzba, Schwilling, Grodd, Krageloh-Mann, & Wilke, 2011; Schmithorst, Holland, & Plante, 2006; Wilke et al., 2005). The Alien Game is designed to involve auditory processing, word retrieval, and the same articulatory processes as in Listen and Name. As such, I hypothesise a contrast of Listen and Name > Alien Game should localise regions supporting semantic access and association (STS/MTG/ITG, left ventral occipito-temporal cortex, Pars Triangularis and Orbitalis, SFG and bilateral AG). If participants

visualise the item being described, this contrast may also identify regions associated with visual imagery (V1 and left occipito-temporal cortex). Activation may also be seen more bilaterally in a sample of children and adolescents (Berl et al., 2014; Brauer & Friederici, 2007), compared to predictions made by the functional neuroanatomical model (Price, 2012).

3.2.4 Read and Name: Sentence-level semantic and syntactic processing from orthography

According to the functional neuroanatomical model (Price, 2012), reading comprehension involves early visual processing (bilateral occipital and left ventral occipito-temporal cortex). At this stage a word may be recognised as a whole visual form and gain direct access to semantics via left posterior MTG and occipito-temporal regions, followed by semantic association and retrieval (left anterior STS/MTG/ITG, left SFG, Pars Orbitalis and Triangularis, bilateral AG, and the temporal poles). Alternatively, the letters can be recoded into phonology via the posterior vOT, Pars Opercularis and precentral gyrus, which is supported by short-term memory (left Pars Opercularis and ventral SMG). The relevant articulatory associations and sequences (bilateral premotor cortex, left Pars Opercularis, and pre-SMA) then provide access to semantics (via posterior STS), followed by semantic association and retrieval (regions listed above). The appropriate response can then be selected (left MFG), along with the articulatory associations (regions listed above), before motor output (SMA, ACC, bilateral precentral gyri, left putamen, insula and thalamus) and auditory feedback (bilateral STG and left posterior STS).

The Read and Name game (RN) involves all of the above processing stages outlined in the functional neuroanatomical model, while Face Finder (FF) involves only early visual processing, word retrieval, articulatory associations and motor output. As such, I hypothesise a contrast of RN > FF to show activation in regions associated with sematic access via either the direct or indirect route, and in regions associated with semantic processing. In adults these include: left posterior MTG, ventral occipito-temporal cortex, Pars Opercularis, precentral gyrus, left ventral SMG, bilateral premotor cortex, pre-SMA and posterior STS,

left anterior STS/MTG/ITG, left SFG, Pars Orbitalis and Triangularis, bilateral AG, and the temporal poles. Children engage similar regions during reading comprehension, with no evidence for developmental change in core comprehension cortex (regions associated with sentence-level semantic and syntactic processing) (Berl et al., 2010). Relative to adults however, children may be expected to show reduced activation in left MTG and fusiform gyrus (Booth et al., 2001; Turkeltaub, Gareau, Flowers, Zeffiro, & Eden, 2003); regions associated with direct semantic access via the whole visual word form in the functional neuroanatomical model (Price, 2012). Conversely, children may show additional activation in the right ITS and fusiform gyrus (Turkeltaub et al., 2003). It has been proposed that activation of these right inferior temporal regions in early learners reflects a non-lexical 'form recognition' strategy for identifying words, which becomes disengaged as reading via phonological routes improves (Orton, 1935; Turkeltaub et al., 2003).

3.3 Amodal semantic processing regions

The Panda Games battery has been designed to engage different functional sub-networks within the language system, to provide flexible mapping in pre-surgical candidates. Having introduced my hypotheses for the specific networks supporting each task, I now outline hypotheses regarding activation associated with semantic processing common to all tasks. This analysis aims to identify a 'core' semantic network in children, which may form the back-bone of the language system, and may be particularly crucial for language development and post-surgical language outcome.

The functional neuroanatomical model predicts common activation across tasks associated with; 1) semantic access in left posterior MTG/ITG, 2) semantic associations in left anterior temporal cortex (including anterior STS, MTG, ITG and vOT), 3) semantic retrieval in fronto-parietal cortex (including left SFG, left Pars Triangularis and Orbitalis, and bilateral angular gyri), and 4) multimodal integration in the posterior STS and temporal poles. Other reviews of the functional imaging literature specifically aimed at identifying regions associated with amodal semantic

processing propose a similar left lateralised network including; posterior IPL, MTG, vOT (and bordering parahippocampal gyri), medial SFG, IFG, ventromedial prefrontal cortex, and PCC (Binder, 2009).

Studies of patients with brain lesions have further contributed to our understanding of this network, by identifying which of these regions appear *critical* to semantic processing (at least in adults). Studies of patients with semantic dementia suggest the bilateral anterior temporal lobes are critical to semantic processing (Hodges & Patterson, 2007; Hodges et al., 1992; Jefferies, Patterson, & Ralph, 2008; Mummery et al., 2000). Specifically, these regions are critical for storing amodal semantic representations of an item (information from general knowledge and personal experience, and sensory-motor associations, which together form the concept of an item). Mappings between the location of structural damage and different types of naming errors in patients with left hemisphere stroke (DeLeon et al., 2007) suggest; 1) the left STG and MTG are critical for lexical-semantic representations (the defining features of a semantic representation, which are associated with a word form), and 2) left vOT and angular gyrus are critical to amodal lexical-semantic access (access to the phonological or orthographic word form).

Because there are no previous investigations of amodal semantic processing in children, I have no prior hypotheses to suggest this network will differ in children and adolescents, relative to adults. Based on functional imaging and patient lesion studies conducted in adults, I hypothesise the core network supporting amodal semantic processing throughout the Panda Games task battery will include; left posterior MTG/ITG, bilateral temporal pole, vOT, left lateral or medial SFG, left Pars Triangularis and Orbitalis, and bilateral angular gyri.

4. METHODS

4.1 Participants

Forty three healthy children performed the Panda Games battery (Chapter 3), including Picture Naming (PN), Picture Describing (PD), Listen and Name (LN) and

Read and Name (RN), and their respective baseline conditions; Colour Naming with words (CNw), Colour Naming with sentences (CNs), Alien Game (AG) and Face Finder (FF). To reduce scanning time for very young children, children aged <8 years did not perform RN or FF (*N*=17). Four participants >8 years of age also failed to perform one or both of these tasks due to visual problems not reported on the inclusion questionnaire.

4.2 Neuropsychological assessment

Non-verbal ability (performance IQ) and executive functioning (General Executive Composite) were assessed according to Chapter 3. Verbal abilities were assessed using the CELF-IV and CTOPP, as reported in Chapter 3. Verbal measures relevant to each task are reported in Table 16- Table 19.

4.3 Task practice and pre-scan preparation

Participants were given the opportunity to practice the Panda Games online prior to their appointment (as described in Chapter 3). All participants went through a '*How to play*' booklet before their scan, which included task instructions and examples (Chapter 3 and Appendix 8.). Online task practice is reported (in minutes). Children <13 years of age and those who were nervous underwent mock scanning (Chapter 3).

4.4 Functional MRI

Functional MR images were acquired and pre-processed as per Chapter 3, with the exception that in-scanner movement was corrected using the motion fingerprint method (Wilke, 2012), as described in Chapter 4. All participants performed the Panda Games (Chapter 3). Only correct trials were analysed, to control for task performance accuracy.

4.5 Participant exclusions (scan failures)

Participants were excluded on a task-by-task basis according to the following criteria 1) maximum in-scanner movement >4 mm (1 voxel plus gap) during either the task or baseline condition, 2) severe movement artefact in the z direction as described in Chapter 6, 3) poor task performance (<60% correct trials) for either the task or

baseline condition, and 4) other factors associated with scan failure or poor scan quality (including visual problems not previously reported, scanner faults, artefact and falling asleep inside the scanner).

4.6 Effect of mock scanning

Participants underwent mock scanning if they were young (aged 5-7 years) or expressed nerves or anxiety. Mock scanning was not possible in all cases, due to time limitations (e.g. during after school appointments). Randomised allocation to the mock and no mock condition was therefore not possible. The effect of mock scanning on a) in-scanner movement (average movement during the whole Panda Games battery), b) task performance (mean % correct across the Panda Games), and c) scan failure rates was investigated in sub-groups matched for age, sex, IQ, practice time and pre-scan anxiety: *mock scan* group (N=13, 9 females, mean age=11 years, SD=3), and *no mock scan* group (N=13, 8 females, mean age=11 years, SD=3).

4.7 Sample characteristics

Sample characteristics for each group map analysis were compared by age and sex using appropriate parametric and non-parametric analyses, to ensure all groups were well-matched on these variables. Practice time and in-scanner movement were also compared between conditions (task and baseline) to identify any potential confounds which may need to be controlled during group map analyses or in later chapters. All of these analyses are reported in Appendix 10.

4.8 Task performance

The percentage of semantically correct trials was used as a measure of task performance, and is reported for each task. The association of task performance with age was investigated between age groups using one-way ANOVA or the Kruksall-Wallis test, post-hoc independent samples t-tests or Man Whitney U tests, and also using correlation analyses (Pearson's r or Spearmen's Rho). The effects of sex and task (task or baseline) on performance accuracy were also investigated using independent samples t-tests or Mann Whitney U tests, and paired samples t-tests or Wilcoxon-Signed Ranks test, respectively. Repeated-measures ANOVA was performed to investigate differences in performance accuracy between the four

language tasks in the Panda Games battery (PN, PD, LN and RN) in participants who performed all four tasks.

4.9 Network activation

For each of the Panda Games, individual participant contrast images were entered into three one-sample t-tests at the second-level, to identify; 1) the overall network (task >rest), including regions which may be associated with executive and sensory-motor aspects of task performance, 2) the network supporting executive and sensory-motor aspects of task performance, with minimal (or no) linguistic (i.e. semantic or syntactic) processing demands (baseline>rest), and 3) regions associated with semantic and syntactic processing (task>baseline).

4.10 Identifying amodal semantic processing regions

Amodal semantic processing was investigated in a sub-group of participants with good quality data who performed all three naming tasks PN, LN and RN (N=18, 9-16 years). Sample descriptives for this analysis are displayed in Table 19. At the second-level I performed conjunction analysis across tasks (PN>CNw, LN>AG and RN>FF) to identify regions of activation associated with semantic processing across all three input modalities; visual (nonverbal), phonological and orthographical. In cases where results did not reach significance at p<0.05 FWE corrected, lower statistical thresholds are reported (p<0.001 and p<0.005) to reduce the rate of false negatives when using these very high-level contrasts. Statistical thresholds are clearly labelled throughout.

5. RESULTS

5.1 Scan failure rates and causes

Scan failure rates across the Panda Games battery are summarised in Figure 34, and are described for each task in the sections below. Movement – as the largest cause of scan failure – was investigated further to identify influential variables which future studies could target to reduce failure rates. Correlation analyses showed mean

movement decreased with age (rho=-.37, *p*= 0.02), as did large sporadic movements (maximum movement; rho=-.36, *p*=0.02). Movement did not differ between males and females, and showed no relation to socio-economic status, overall ability (IQ) or executive functioning (which includes measures of behavioural regulation). Movement was not associated with anxiety during scanning or scores on a measure of researcher interaction (Chapter 3, section 3.7). Movement was associated with scores on measures of *in-scanner experience* (rho=-.43, *p*=0.04) and *head-phone comfort* (rho=-.47, *p*=0.02) (Chapter 3, section 3.7), suggesting participants who had a more negative experience inside the scanner, and for whom the head phones were less comfortable, moved more inside the scanner.

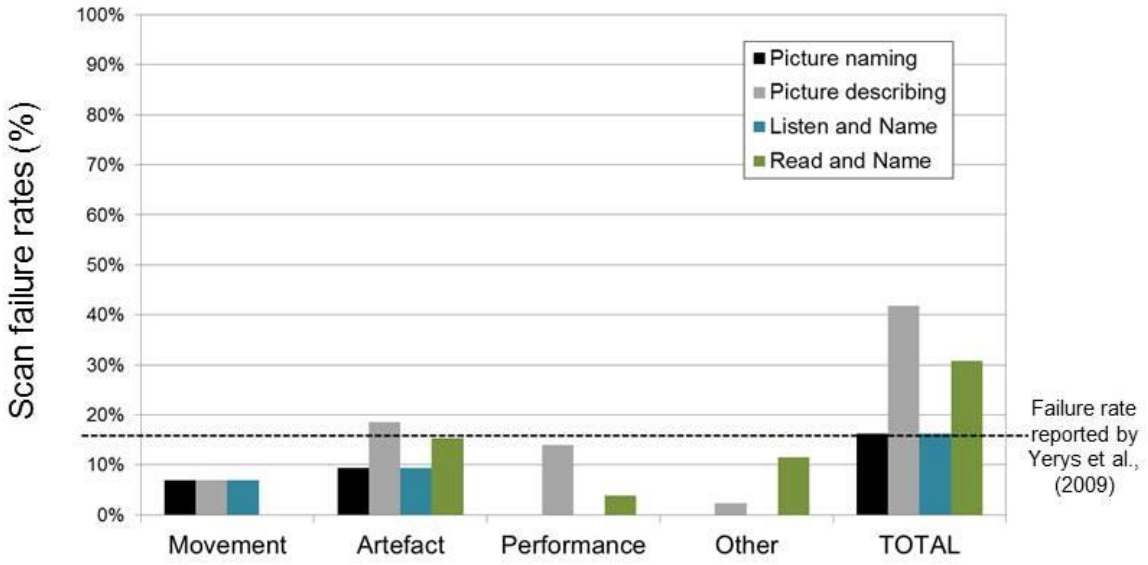


Figure 34: Scan failure rates and causes for the Panda Games in healthy children and adolescents aged 5-16 years. The failure rate reported by Yerys et al., (2009) is marked with a dashed line, as a point of comparison.

5.1.1 Picture naming

Of the 43 participants who performed both PN and CNw, seven were excluded; three due to in-scanner movement and four due to movement-related artefact. Descriptives for the final sample (N=36, 5-16 years) are presented in Table 16.

5.1.2 Picture describing

Of the 43 participants who performed both PD and CNs, 18 were excluded from group map analyses; 1 due to scanner artefact associated with the scanner over-heating, 11 due to in-scanner movement (8 of whom showed movement artefact) and six due to poor task performance accuracy. Of those with poor performance accuracy for PD, one participant found it difficult to perform the task because stimuli turned black and white during scanning (due to a loose connection) but this was not reported at the time, four had visual difficulties not reported on the inclusion questionnaire, and 1 had fallen asleep during the previous scan and was disorientated/sleepy during performance of PD. Two participants gave partial responses (a subject-verb structure e.g. "The squirrel is jumping"). As these responses were semantically correct and consistent across trials, and as activation was similar at the first-level in comparison to other participants, I included these participants in the group map analysis (classing partial responses as correct). Descriptives for the final sample (N=25, 7-16 years) are presented in Table 17.

5.1.3 Listen and Name

Of the 43 participants who performed both LN and AG, 7 participants were excluded from group map analyses due to in-scanner movement during either the task or baseline condition. Descriptives for the final sample (N=36, 5-16 years) are presented in Table 18.

5.1.4 Read and Name

Of the 26 participants who performed both RN and FF during scanning, 8 were excluded from group map analyses; four due to speech-related movement artefact, one due to poor task performance, and three due to poor data quality (lack of activation) despite good in-scanner parameters (good performance and low movement). These three participants with poor data quality were excluded to remove the influence of outliers on group-level results in this relatively small sample. Descriptive statistics for the final sample (N=18, 9-16 years) are displayed in Table 19.

5.2 Effect of mock scanning

Results are displayed in Table 15. There was no significant effect of mock scan on in-scanner movement or scan failure rates. Task performance accuracy was improved for children who underwent mock scanning ($t(24)=-2.69$, $p<0.01$), despite equal practice time between groups.

	No mock		Mock		
Performance measures	**Mean**	**SD**	**Mean**	**SD**	***p***
Mean in-scanner movement (*mm*)	0.15	0.07	0.13	0.06	*0.50*
Mean task performance (*% correct*)	89%	5%	93%	2%	*0.01*
Scan failure rate					
Picture Naming	8%	-	15%	-	*0.52*
Picture Describing	31%	-	23%	-	*0.47*
Listen and Name	8%	-	15%	-	*0.52*
Read and Name	23%	-	8%	-	*0.80*

Table 15: The effect of mock scanning on scan quality in healthy children. The mean and standard deviation (SD) for in-scanner movement, task performance accuracy and failure rates are shown for children who underwent mock scanning (mock, N=13) and did not undergo mock scanning (No mock, N=13), matched for age, sex, IQ and anxiety. The results of group comparisons are shown (*p*; black=significant).

5.3 Group-map sample characteristics

Demographic, neuropsychological, task practice and movement data are displayed for participants included in group map analyses for each of the Panda Games (Table 16 - Table 19). Overall, age groups were well-matched for demographics, neuropsychological performance, task practice time, task performance accuracy and in-scanner movement for each task. Language tasks and their baseline condition were also well-matched for practice time and in-scanner movement. Elaboration of significant or trend-level age and task effects, as well as results from between-sex comparisons, are reported for each task in Appendix 10.

	Total sample (5-16 years)			Early childhood (5-7 years)			Late childhood (9-12 years)			Adolescence (13-16 years)		
	N	**Mean**	SD	*N*	**Mean**	SD	*N*	**Mean**	SD	*N*	**Mean**	SD
Demographics												
Age *(years)*	*36*	**11.37**	3.39	*11*	**7.37**	1.05	*12*	**11.02**	1.05	*13*	**15.08**	1.34
Sex *(female:male)*	*36*	**19:17**	-	*11*	**5:6**	-	*12*	**9:3**	-	*13*	**5:8**	-
Handedness *(typical:atypical)*	*34**	**75:25**	-	*11*	**0.42**	-	*11*	**10:1**	-	*12*	**8:4**	-
Mother's education *(years)*	*34**	**22.5**	5.00	*11*	**22.27**	1.74	*11*	**21.82**	3.82	*12*	**23.33**	7.60
Neuropsychological performance												
Verbal *(standard scores)*												
Core language	*30**	**111**	11	*8*	**109**	9	*11*	**114**	15	*11*	**109**	9
Formulated Sentences	*34**	**13**	8	*11*	**12**	4	*11*	**12**	3	*12*	**15**	13
Non-verbal *(standard scores)*												
Performance IQ	*34**	**118**	17	*11*	**122**	21	*11*	**121**	16	*12*	**111**	12
General executive composite	*34**	**49**	9	*11*	**51**	10	*11*	**47**	5	*12*	**50**	10
Pre-scan task practice												
Picture Naming *(minutes)*	*36*	**2.12**	2.35	*11*	**2.27**	2.29	*12*	**2.72**	2.53	*13*	**1.46**	2.24
Colour Naming *(minutes)*	*36*	**0.99**	1.49	*11*	**1.08**	1.37	*12*	**1.41**	1.75	*13*	**0.52**	1.30
In-scanner performance												
Task performance *(% correct)*												
Picture Naming	*36*	**97%**	7%	*11*	**98%**	2%	*12*	**100%**	0%	*13*	**93%**	10%
Colour Naming	*36*	**100%**	1%	*11*	**100%**	0%	*12*	**100%**	1%	*13*	**100%**	0%
Mean movement *(mm/TR)*												
Picture Naming	*36*	**0.12**	0.08	*11*	**0.14**	0.13	*12*	**0.11**	0.05	*13*	**0.12**	0.06
Colour Naming	*36*	**0.12**	0.09	*11*	**0.13**	0.09	*12*	**0.12**	0.13	*13*	**0.11**	0.04
Maximum movement *(mm/TR)*												
Picture Naming	*36*	**1.02**	0.67	*11*	**1.15**	0.86	*12*	**0.99**	0.56	*13*	**0.93**	0.62
Colour Naming	*36*	**1.05**	0.59	*11*	**1.31**	0.75	*12*	**0.83**	0.51	*13*	**1.04**	0.46

Table 16: Descriptive statistics for participants included in the group map analysis of Picture Naming. Sample size (*N*), mean and standard deviation (SD) are provided for all measures. Missing neuropsychological data are indicated (*) and are due to failure to return for neuropsychological assessment (n=2) or fatigue.

	Total sample (7-16 years)			Early childhood (7-10 years)			Late childhood (11-12 years)			Adolescence (13-16 years)		
	N	**Mean**	SD	*N*	**Mean**	SD	*N*	**Mean**	SD	*N*	**Mean**	SD
Demographics												
Age *(years)*	*25*	**11.45**	2.93	*10*	**8.55**	1.27	*8*	**11.79**	0.64	*7*	**15.18**	1.00
Sex *(female:male)*	*25*	**12:13**	-	*10*	**5:5**	-	*8*	**5:3**	-	*7*	**2:5**	-
Handedness *(typical:atypical)*	*24**	**22:2**	-	*10*	**9:1**	-	*7*	**7:0**	-	*7*	**6:1**	-
Mother's education *(years)*	*24**	**22.42**	4.88	*10*	**23.40**	2.80	*7*	**23**	7.53	*7*	**20.43**	4.04
Neuropsychological performance												
Verbal *(standard scores)*												
Expressive language	*23**	**111**	13	*9*	**116**	14	*7*	**109**	13	*7*	**107**	10
Formulated Sentences	*24**	**12**	3	*10*	**12**	3	*7*	**12**	3	*7*	**11**	2
Non-verbal *(standard scores)*												
Performance IQ	*24**	**120**	17	*10*	**131**	15	*7*	**116**	13	*7*	**108**	13
General executive composite	*24**	**49**	9	*10*	**45**	8	*7*	**48**	6	*7*	**55**	9
Pre-scan task practice												
Picture Describing *(minutes)*	*25*	**1.06**	2.52	*10*	**1.94**	3.84	*8*	**0.67**	0.73	*7*	**0.24**	0.64
Colour Naming (w/sentences) *(minutes)*	*25*	**0.24**	0.63	*10*	**0.46**	0.86	*8*	**0.18**	0.52	*7*	**0.00**	0.00
In-scanner performance												
Task performance *(% correct)*												
Picture Describing	*25*	**77%**	21%	*10*	**83%**	13%	*8*	**70%**	28%	*7*	**76%**	23%
Colour Naming (w/sentences)	*25*	**99%**	5%	*10*	**97%**	8%	*8*	**100%**	0%	*7*	**100%**	0%
Mean movement *(mm/TR)*												
Picture Describing	*25*	**0.17**	0.10	*10*	**0.15**	0.08	*8*	**0.16**	0.07	*7*	**0.21**	0.15
Colour Naming (w/sentences)	*25*	**0.13**	0.07	*10*	**0.14**	0.07	*8*	**0.10**	0.03	*7*	**0.16**	0.11
Maximum movement *(mm/TR)*												
Picture Describing	*25*	**1.47**	1.02	*10*	**1.43**	0.94	*8*	**1.38**	0.72	*7*	**1.64**	1.47
Colour Naming (w/sentences)	*25*	**1.00**	0.47	*10*	**1.15**	0.59	*8*	**0.86**	0.25	*7*	**0.97**	0.47

Table 17: Descriptive statistics for participants included in the group map analysis of Picture Describing. Sample size (*N*), mean and standard deviation (SD) are provided for all measures. Missing neuropsychological data are indicated (*) and are due to failure to return for neuropsychological assessment (n=2) or fatigue.

	Total sample (5-16 years)			Early childhood (5-8 years)			Late childhood (9-12 years)			Adolescent (13-16 years)		
	N	**Mean**	SD	*N*	**Mean**	SD	*N*	**Mean**	SD	*N*	**Mean**	SD
Demographics												
Age *(years)*	*36*	**10.96**	3.27	*11*	**7.09**	0.962	*12*	**10.6**	1.003	*13*	**14.57**	1.229
Sex *(female:male)*	*36*	**19:17**	-	*11*	**5:6**	-	*12*	**9:3**	-	*13*	**5:8**	-
Handedness *(typical:atypical)*	*34**	**28:6**	-	*11*	**10:1**	-	*11*	**10:1**	-	*12*	**8:4**	-
Mother's education *(years)*	*34**	**22.71**	5.048	*11*	**22.91**	2.166	*11*	**21.82**	3.816	*12*	**23.33**	7.596
Neuropsychological performance												
Verbal *(scaled scores)*												
Receptive language	*34**	**12**	3	*11*	**12**	2	*11*	**13**	3	*12*	**11**	3
Non-verbal *(standard scores)*												
Performance IQ	*33**	**120**	17	*10*	**129**	18	*11*	**121**	16	*12*	**111**	12
General executive composite	*34**	**49**	9	*11*	**51**	10	*11*	**47**	5	*12*	**50**	10
Pre-scan task practice												
Listen and Name *(minutes)*	*36*	**1.83**	3.43	*11*	**2.72**	3.80	*12*	**2.18**	4.40	*13*	**0.76**	1.60
Alien Game *(minutes)*	*36*	**1.77**	2.31	*11*	**2.56**	2.19	*12*	**1.89**	2.80	*13*	**0.99**	1.78
In-scanner performance												
Task performance *(% correct)*												
Listen and Name	*36*	**87%**	11%	*11*	**86%**	11%	*12*	**92%**	7%	*13*	**84%**	12%
Alien Game	*36*	**88%**	10%	*11*	**83%**	11%	*12*	**90%**	10%	*13*	**91%**	9%
Mean movement *(mm/TR)*												
Listen and Name	*36*	**0.12**	0.06	*11*	**0.13**	0.05	*12*	**0.12**	0.07	*13*	**0.11**	0.05
Alien Game	*36*	**0.13**	0.08	*11*	**0.15**	0.07	*12*	**0.11**	0.05	*13*	**0.15**	0.10
Maximum movement *(mm/TR)*												
Listen and Name	*36*	**1.11**	0.76	*11*	**1.44**	0.88	*12*	**1.01**	0.89	*13*	**0.93**	0.42
Alien Game	*36*	**1.19**	0.80	*11*	**1.58**	0.91	*12*	**1.02**	0.81	*13*	**1.01**	0.59

Table 18: Descriptive statistics for participants included in the group map analysis of Listen and Name. Sample size (N), mean and standard deviation (SD) are provided for all measures. Missing neuropsychological data are indicated (*) and are due to failure to return for neuropsychological assessment (n=2) or fatigue (n=2-3).

	Total sample (9-16 years)			Late childhood (9-12 years)			Adolescence (13-16 years)		
	N	**Mean**	SD	*N*	**Mean**	SD	*N*	**Mean**	SD
Demographics									
Age *(years)*	*18*	**13.56**	2.38	*8*	**11.24**	1.14	*10*	**15.41**	1.07
Sex *(female:male)*	*18*	**9:9**	-	*8*	**5:3**	-	*10*	**4:6**	-
Handedness *(typical:atypical)*	*16**	**12:4**	-	*7*	**6:1**	-	*9*	**6:3**	-
Mother's education *(years)*	*16**	**22.63**	5.55	*7*	**22.43**	3.91	*9*	**22.78**	6.80
Neuropsychological performance									
Verbal *(standard scores)*									
Core language	*15*	**111**	11	*7*	**112**	14	*8*	**110**	9
Phonological Awareness	*16**	**99**	9	*7*	**105**	7	*9*	**95**	8
Rapid Naming	*16**	**102**	10	*7*	**103**	10	*9*	**102**	10
Non-verbal *(standard scores)*									
Performance IQ	*16**	**114**	14	*7*	**120**	15	*9*	**109**	12
General executive composite	*16**	**50**	7	*7*	**48**	5	*9*	**51**	8
Pre-scan task practice									
Read and Name *(minutes)*	*18*	**0.78**	1.26	*8*	**0.87**	1.31	*10*	**0.72**	1.29
Face Finder *(minutes)*	*18*	**0.74**	1.05	*8*	**0.97**	1.10	*10*	**0.55**	1.02
In-scanner performance									
Task performance *(% correct)*									
Read and Name	*18*	**89%**	12%	*8*	**93%**	7%	*10*	**85%**	13%
Face Finder	*18*	**88%**	11%	*8*	**87%**	12%	*10*	**90%**	11%
Mean movement *(mm/TR)*									
Read and Name	*18*	**0.11**	0.08	*8*	**0.13**	0.11	*10*	**0.11**	0.04
Face Finder	*18*	**0.13**	0.07	*8*	**0.13**	0.07	*10*	**0.14**	0.07
Maximum movement *(mm/TR)*									
Read and Name	*18*	**0.92**	0.54	*8*	**1.08**	0.73	*10*	**0.80**	0.31
Face Finder	*18*	**1.17**	0.88	*8*	**1.40**	1.13	*10*	**0.99**	0.62

Table 19: Descriptive statistics for participants included in the group map analysis of Read and Name. Sample size (N), mean and standard deviation (SD) are provided for all measures. Missing neuropsychological data are indicated (*) and are due to failure to return for neuropsychological assessment (n=2) or fatigue.

5.4 In-scanner task performance

5.4.1 Age

Task performance accuracy is shown across age for all of the Panda Games in Figure 35.

5.4.1.1 Picture naming

Kruskal-Wallis test showed a significant difference in PN performance between age groups ($X^2(2)=15.22$, $p<0.001$). Post hoc Mann-Whitney U tests confirmed participants in later childhood performed better than participants in early childhood ($U=30$, $p=0.03$) and adolescents ($U=0.18$, $p=0.001$), suggesting this was not a linear age-effect. Encouragingly, there was no significant correlation of PN performance accuracy and age ($p=0.13$). There were no significant differences in CNw performance by age group ($p=0.37$), and no correlation of CNw performance with age ($p=0.13$).

5.4.1.2 Picture describing

There was no significant difference in PD task performance between age groups ($p=0.45$), and no correlation of PD performance with age ($p=0.37$). There was a trend suggesting participants in early childhood performed CNs worse compared to older participants ($p=0.09$). There was also a significant correlation suggesting performance accuracy for CNs improved with age (rho=.55, $p=0.005$). This correlation was driven by three participants aged 7 years; one initially forgot the task instructions and responded with single colour names (as per CNw) in the first block of stimuli (therefore achieving an accuracy of 75%), and two participants achieved an accuracy of 96% (1 error/omission). All other participants performed with 100% accuracy.

5.4.1.3 Listen and Name

There was no significant difference in LN performance between age groups ($p=0.17$) and correlation analyses showed no significant relationship between age and LN performance ($p=0.34$). Kruskal-Wallis test showed a trend for between-group differences in AG performance accuracy however ($X^2(2)=5.14$, $p=0.08$). Post-hoc Mann-Whitney U tests showed performance was better in

late childhood (trend; p=0.08) and adolescence (U=36, p=0.04), compared to early childhood. There was also a significant correlation of AG performance accuracy with age (rho=0.36, p=0.03), suggesting older participants found it easier to accurately distinguish the sex of the speaker when sounds were spectrally rotated. Younger children still made as many guesses as older children however (the rate of omissions did not correlate with age; p=0.19) suggesting they engaged with the task well, but failed to perform it as accurately compared to older children.

5.4.1.4 Read and Name

There was no significant difference in task performance for RN or FF between age groups (p values>0.18), and there was no significant correlation of task performance with age (p values>0.15).

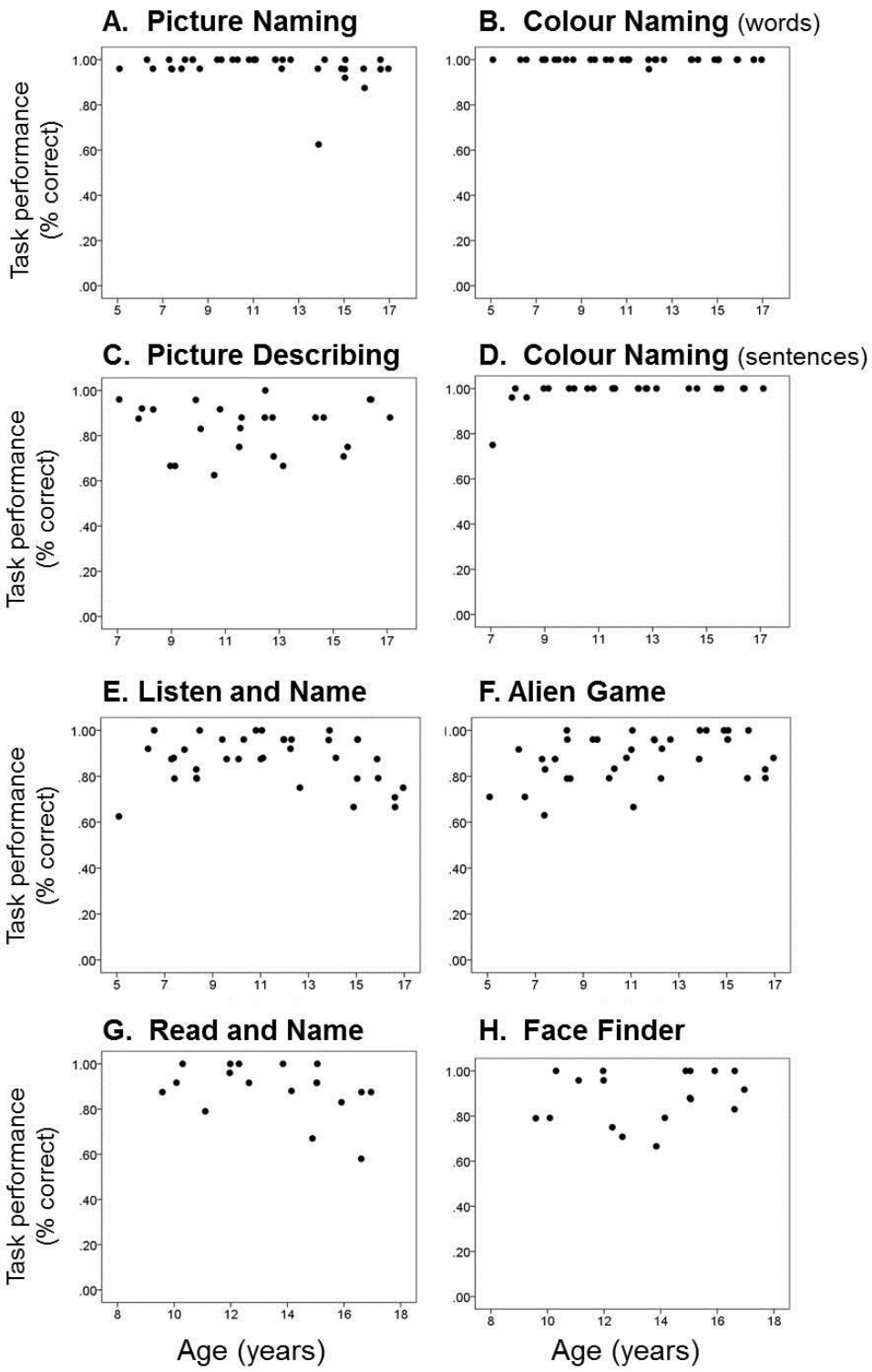


Figure 35: In-scanner task performance accuracy across age in healthy children, for all tasks within the Panda Games battery (A-H). Task performance accuracy (% correct) is shown for participants included in group map analyses for each of the Panda Games: Picture Naming and Colour Naming with words (N=36, 5-16 years), Picture Describing and Colour Naming with sentences (N=25, 7-16 years), Listen and Name

and Alien Game (N=36, 5-7 years), and Read and Name and Face Finder (N=18, 9-16 years).

5.4.1.5 Whole sample analyses of age and task performance

To ensure exclusion of participants with poor performance from group map analyses did not render analysis of age-related trends in task performance (section 5.4.1) invalid, correlation analyses were re-calculated including all participants who performed the task, regardless of exclusion criteria (N=43 for PN, PD, LN, CNw, CNs and AG; N=27 for RN; N=26 for FF). This analysis included participants who had been excluded in sections 5.4.1.1 - 5.4.1.4 due to poor task performance (<60% accuracy). Encouragingly, this analysis failed to show any correlation of age with performance accuracy for PN, PD, LN, RN (although trend; rho=-.36, p=0.07), CNw or FF. Age was significantly correlated with CNs (rho=0.44, p=0.003) and AG (rho=.52, p<0.001), suggesting performance on these tasks improved with age. For CNs, all participants performed the task with >60% accuracy, however only children aged 5-7 years performed below 100% accuracy, achieving: 75% (N=1), 92-96% (N=6) or 100% (N=5) accuracy. For AG, there was a significant difference in task performance between age groups ($X^2(2)=9.63$, p=0.008), with post hoc Mann Whitney U tests showing participants aged 5-8 years performed significantly worse (mean=76% correct, SD=17%) than participants aged 9-12 years (U=52, p=0.02; mean=90% correct, SD=10%) and 13-16 years (U=49, p=0.006; mean=91%, SD=9%). The percentage of all participants who performed <60% and >60% are displayed in Figure 36, as well as mean performance accuracy across the sample for each task.

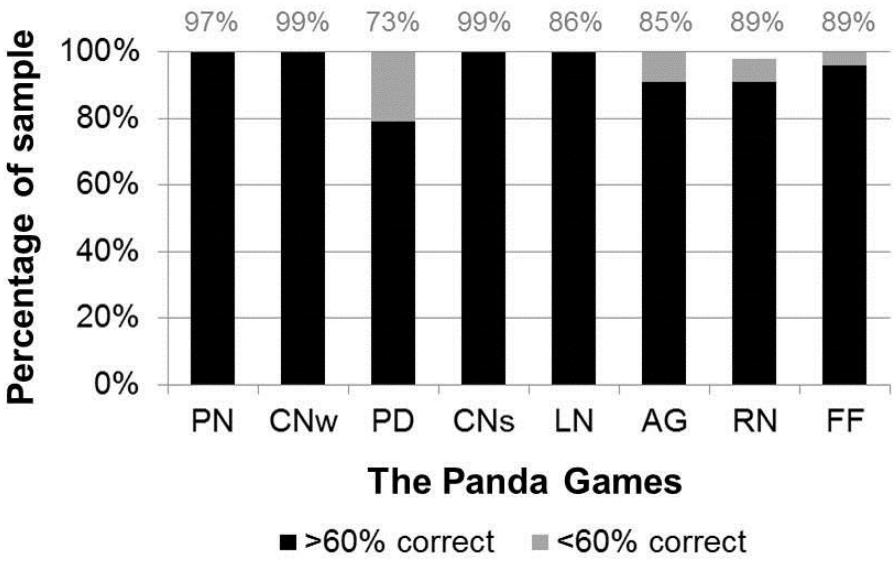


Figure 36: Percentage of participants who achieved good performance accuracy (black) or poor performance accuracy (grey) across the entire sample, for each of the Panda Games. Mean task performance (% correct) is displayed for each task at the top of the relevant bar in grey. PN = Picture Naming, CNw = Colour Naming with words, PD = Picture Describing, CNs = Colour Naming with sentences, LN = Listen and Name, AG = Alien Game, RN = Read and Name, FF = Face Finder.

5.4.2 Sex

There was no significant difference in task performance between males and females for PN or CNw (p values>0.34), PD or CNs (p values >0.44), LN (p=0.98), or RN or FF (p values >0.66). There was a significant difference in AG performance accuracy between males and females (U=98, p=0.04), suggesting females performed the task more accurately (mean=92%, SD=8%) than males (mean=84%, SD=11%). There was no significant interaction of age and sex on AG performance accuracy (p=0.64).

5.4.3 Task (versus baseline)

There was no significant difference in performance accuracy between LN and AG (p=0.38) or between RN and FF (p=0.87). However, participants performed PN

less accurately than CNw (Z=-3.12, p=0.002; Table 16), and PD less accurately than CNs (Z=-3.98, p<0.001; Table 17). The majority of participants performed at ceiling for CNw and CNs (Figure 36).

5.4.4 Language task (PN, PD, LN and RN)

Eighteen participants adhered to group map inclusion criteria for every task in the Panda Games battery (those included in the Read and Name group map; Table 19). Repeated measures ANOVA showed a significant main effect of task on performance accuracy for these participants (F(4)=7.64, p<0.001). Post-hoc paired-samples t tests showed significantly higher performance accuracy for PN (mean=98%, SD=4%) compared to all other tasks: PD (mean=72%, SD=28%; p=0.002), LN (mean=85%, SD=11%; p<0.001), and RN (mean=89%, SD=12%; p=0.005). Task performance was lower for PD versus all other tasks: PN (p=0.002), RN (p=0.046) and at trend level versus LN (p=0.09).

5.5 Network activation

Group map activation results are shown in Figure 37 (PN), Figure 39 (PD), Figure 40 (LN) and Figure 41 (RN). Peak activation for each contrast are listed in Appendix 11. Activation patterns are summarised below.

5.5.1 Picture naming

5.5.1.1 Picture Naming (>rest)

Picture naming activated a ventral network which included visual areas (bilateral posterior inferior occipital gyri), the left ventral occipito-temporal cortex (vOT) and an anterior fusiform/parahippocampal gyrus cluster which incorporated the temporal pole (p<0.05, FWE corrected). At a lower threshold (p<0.001, k>10, uncorrected) bilateral activation was seen extending from inferior occipital regions along the fusiform gyrus and hippocampus, to the temporal poles (Figure 37A). Additional dorsal activation was seen in bilateral precentral gyri, left premotor cortex and the SMA.

5.5.1.2 Colour naming (>rest)

Colour naming was associated with activation in bilateral visual cortex (posterior inferior/middle occipital gyri), right vOT, left posterior lingual gyrus, bilateral precentral gyri and the right hippocampus ($p<0.05$, FWE corrected). At a lower threshold ($p<0.001$, $k>10$, uncorrected) additional activation was seen in the right SMA and left anterior fusiform gyrus (close to the head of the hippocampus) (Figure 37B).

5.5.1.3 Semantic processing from visual input (PN > CNw)

Compared to colour naming, naming a picture was associated with increased activation of bilateral visual regions (bilateral middle occipital gyri and right inferior occipital gyrus), bilateral vOT and left posterior ITG ($p<0.05$, FWE corrected). At a lower threshold additional activation was seen in the left hemisphere, including; amygdala, temporal pole, Pars Triangularis and a small cluster just medial to Pars Orbitalis. Right sided activation was also seen in the anterior parahippocampal gyrus (Figure 37C).

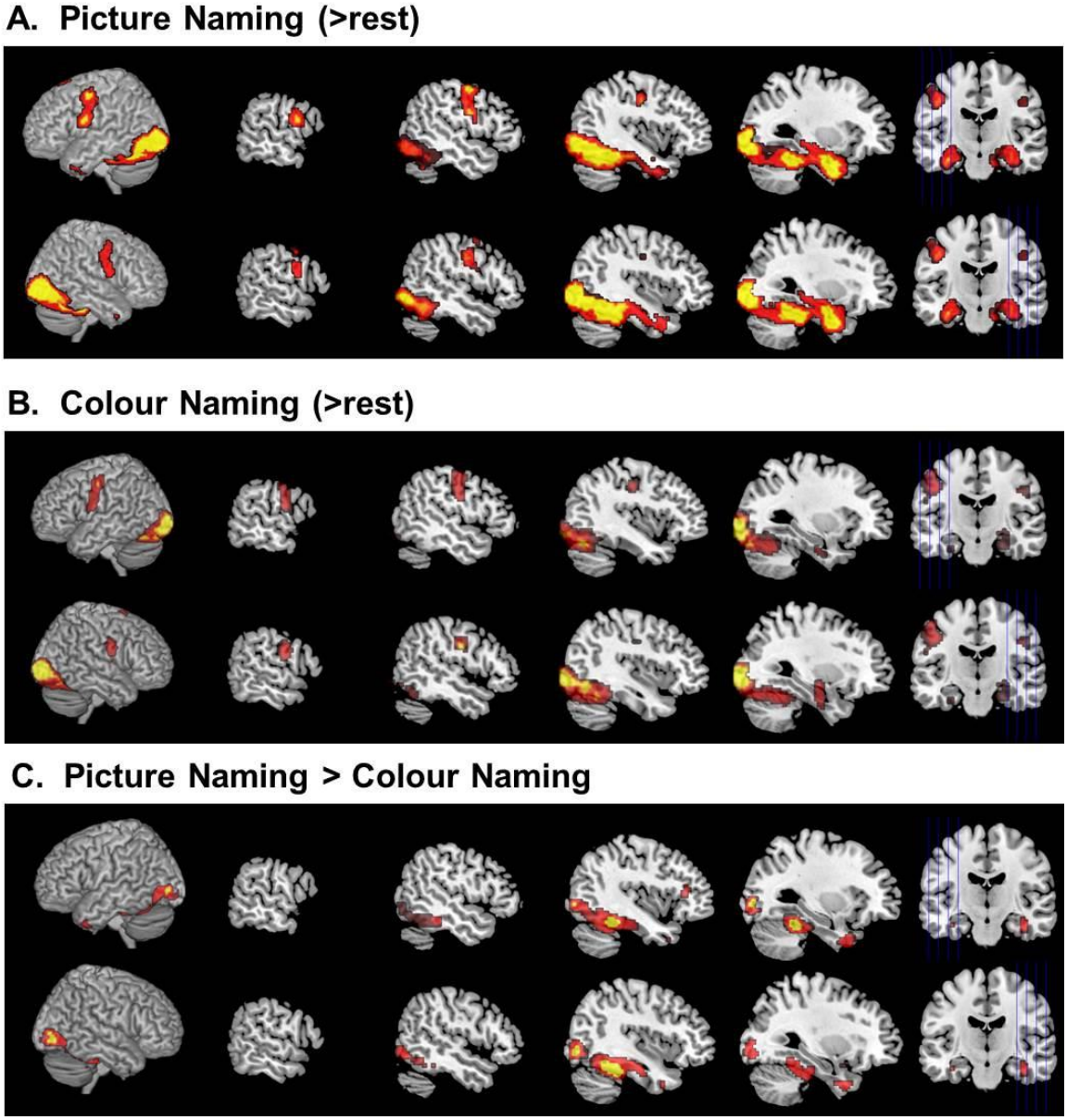


Figure 37: Group activation (N=36, 5-16 years) for word retrieval from pictures (A), colour naming (B) and for semantic processing during word retrieval from pictures (C). For each contrast activation is shown for the left (top) and right (bottom) hemispheres, on a rendered image of the brain (left) and on five sagittal slices (right). Activation is shown at $p<0.05$ FWE corrected (yellow) and $p<0.001$ uncorrected (red).

5.5.2 Picture describing

5.5.2.1 Picture describing (>rest)

Sentence generation was associated with activation in bilateral precentral gyrus, vOT and middle occipital gyrus, right SMA, and left temporal pole and anterior fusiform gyrus (p<0.05, FWE corrected). At a lower threshold, more extensive activation was seen in the SMA and along the length of the fusiform gyri bilaterally (p<0.001, k>10, uncorrected). Contrary to hypotheses, there was no significant activation in Broca's or Wernicke's areas, or in Geschwind's territory (the IPL) (Figure 39A). However, activation was seen in bilateral Pars Triangularis (BA 45) and left Pars Orbitalis (BA 47) at a very low threshold (p<0.01, k>10), suggesting activation of Broca's area was variable across the sample (Figure 38). No activation was seen in Wernicke's area or Geschwind's territory, even when using extremely liberal thresholds (Figure 38), suggesting these regions were not engaged during sentence generation.

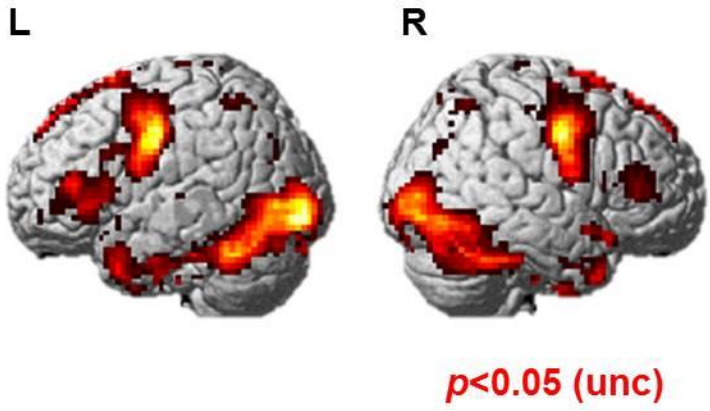


Figure 38: Group map activation (N=25) during sentence generation, at a very liberal statistical threshold. Results are shown for the contrast PD>CNs at p<0.05 uncorrected for multiple comparisons. At this threshold activation is seen in Broca's area, but not Wernicke's area or the inferior parietal lobe.

5.5.2.2 Colour naming with sentences (> rest)

Naming the colour of a picture using a standard sentence structure engaged a similar network to sentence generation (Figure 39B), involving bilateral precentral gyri, SMA and middle occipital gyrus. However, unlike sentence generation there were several peaks of activation within the right hippocampus (p<0.001, k>10, uncorrected), and no activation in vOT or the temporal pole.

5.5.2.3 Semantic and syntactic processing during sentence generation (PD > CNs)

Semantic and syntactic processing during sentence generation was associated with activation in bilateral vOT and left middle occipital gyrus (p<0.001, k>10, uncorrected) (Figure 39C).

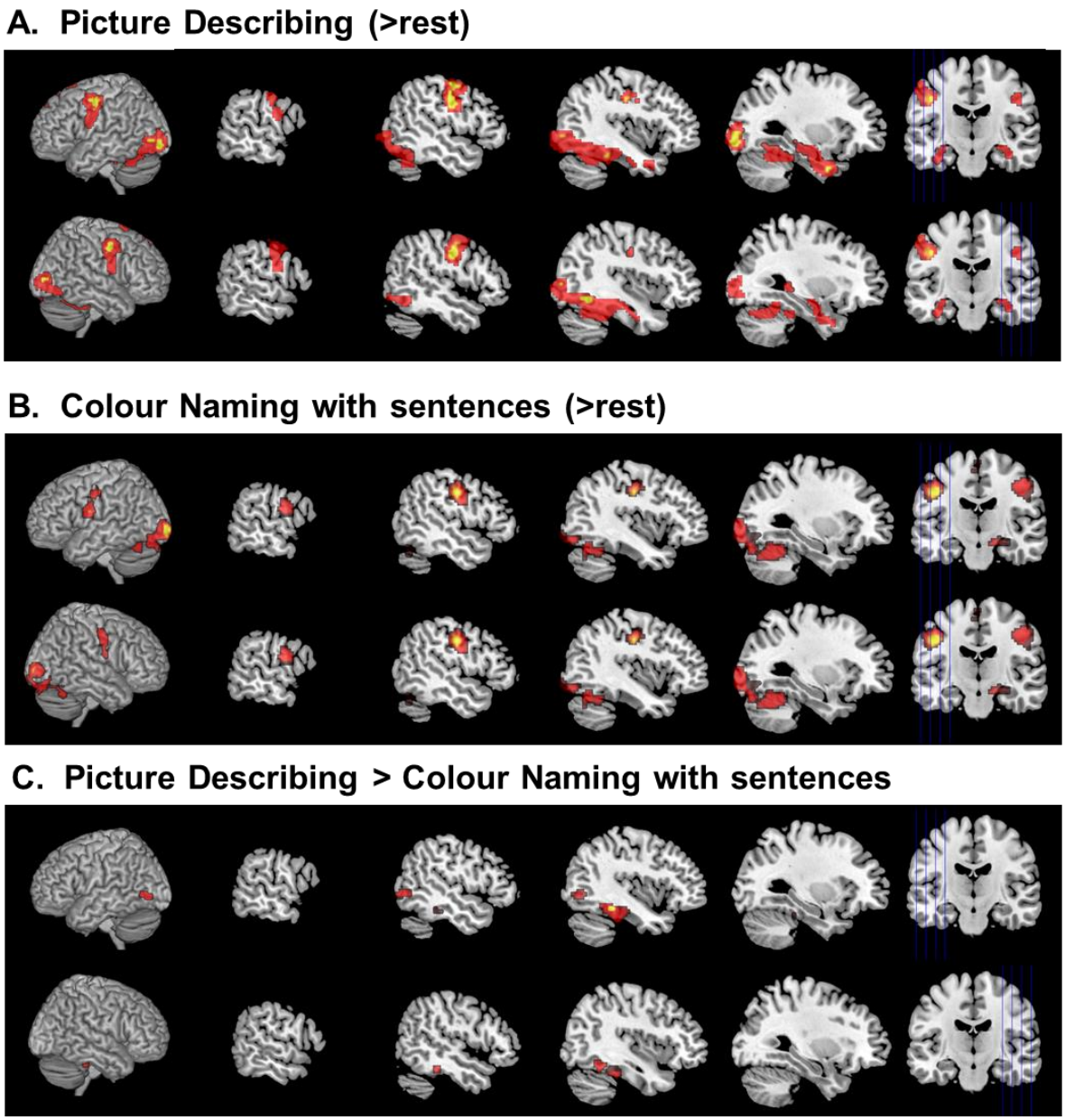


Figure 39: Group activation (N=25, 7-16 years) for sentence generation from pictures (A), for colour naming using a standard sentence structure (B) and for semantic processing during sentence generation from pictures (C). For each contrast activation is shown for the left (top) and right (bottom) hemispheres, on a rendered image of the brain (left) and on five sagittal slices (right). Activation is shown at $p<0.05$ FWE corrected (yellow) and $p<0.001$ uncorrected (red).

5.5.3 Listen and Name

5.5.3.1 Listen and Name (>rest))

On the group spmT map (Figure 40A) significant activation (p<0.05, FWE corrected) was seen bilaterally along the STG, the temporal pole, anterior parahippocampal gyrus/amygdala and hippocampus, medial superior frontal gyrus (BA 8), SMA and dorsal precentral gyrus. Left hemisphere activation was seen in Pars Triangularis (BA 45), vOT and the anterior fusiform gyrus, while right hemisphere activation was seen in the tail of the hippocampus/thalamus. At a lower threshold (p<0.001 uncorrected, k>10) more extensive activation was seen bilaterally in the hippocampus, as well as the thalamus and bilateral globus pallidus. There was also significant activation in left Pars Orbitalis (BA 47) and planum temporale.

5.5.3.2 Alien Game (>rest)

Similar to LN, AG induced significant activation (p<0.05, FWE corrected) bilaterally in the STG (although limited to the mid-posterior section surrounding primary auditory cortex) and the right precentral gyrus (Figure 40B). Left hemisphere activation was seen in the planum temporale. At a lower threshold (p<0.001 uncorrected, k>10) additional activation was seen bilaterally in the lateral globus pallidus, right putamen, left anterior insula, bilateral rolandic operculum, and SMA.

5.5.3.3 Semantic and syntactic processing from phonological input (LN>AG)

Significant activation was seen (Figure 40C) in bilateral vOT, bilateral temporal pole, left anterior fusiform gyrus and anterior STG/STG, Pars Orbitalis, Pars Triangularis and Pars Opercularis (a small cluster deep within the inferior frontal sulcus), and left thalamus. Activation was also seen in the right parahippocampal gyrus/amygdala. Similar results were obtained at a lower threshold (p<0.001 uncorrected, k>10), including medial superior frontal gyrus, right anterior STG, and bilateral lingual gyri.

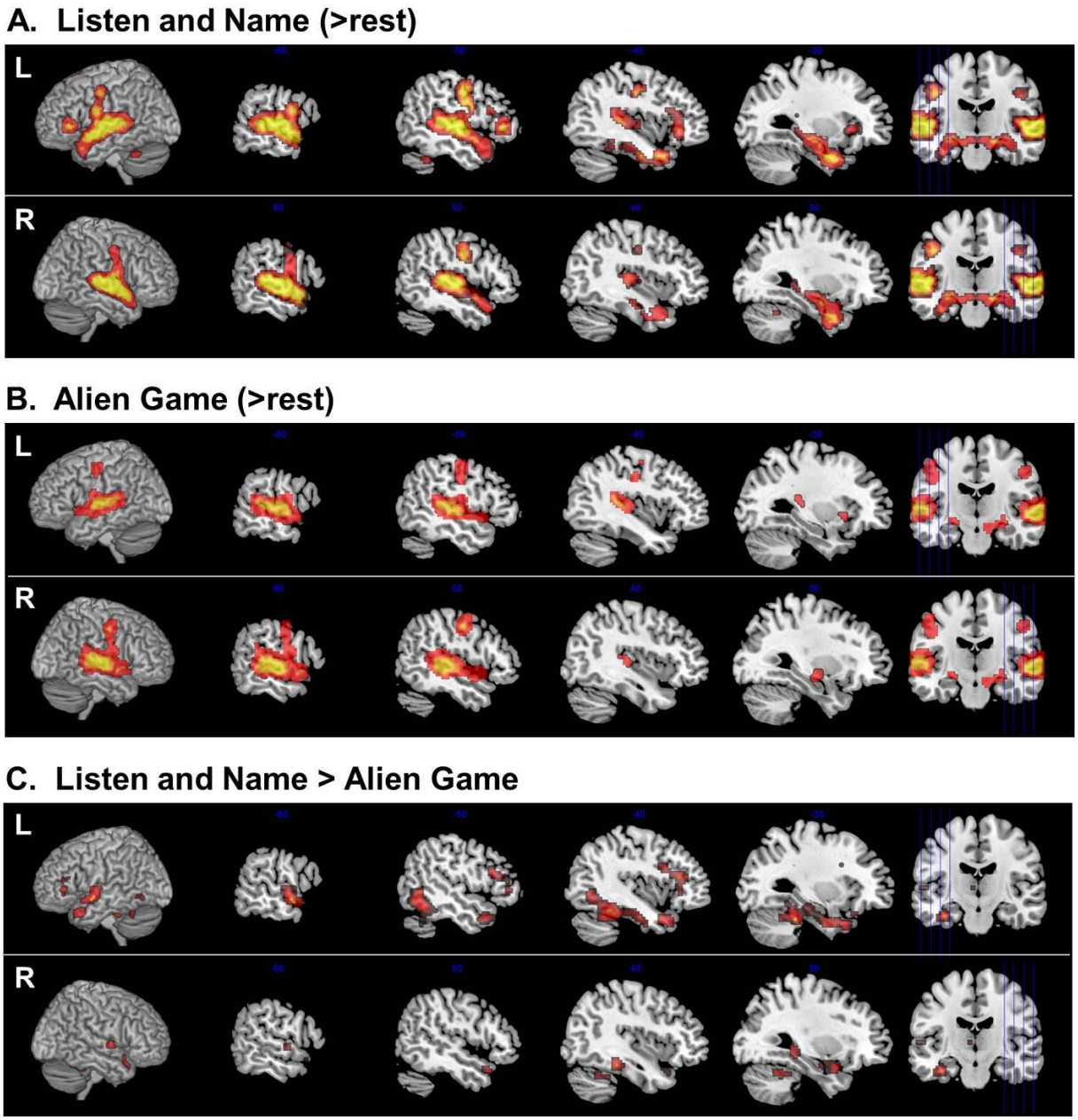


Figure 40: Group activation (N=36, 5-16 years) for auditory comprehension and word retrieval (A), complex auditory processing and word repetition (B) and for semantic and syntactic processing during auditory comprehension (C). For each contrast activation is shown for the left (top) and right (bottom) hemispheres, on a rendered image of the brain (left) and on five sagittal slices (right). Activation is shown at $p<0.05$ FWE corrected (yellow) and $p<0.001$ uncorrected (red).

5.5.4 Read and Name

5.5.4.1 Read and Name (>rest)

Significant activation was found in bilateral lingual gyri in the inferior occipital lobe, left posterior ITG, Pars Triangularis (BA 45) and a dorsal region of the premotor cortex ($p<0.05$, FWE corrected). At a lower threshold ($p<0.001$, $k>10$, uncorrected) additional activation was seen in left posterior MTG, Pars Triangularis (BA 45) and Pars Opercularis (BA 44) (especially deep in the inferior frontal sulcus). Activation in the Pars Triangularis was extensive, and was located within a single medial cluster (at the level of the IFS) which separated into two distinct clusters (anterior and posterior) at the surface of the cortex. Activation was also seen in left vOT and along the length of the left fusiform gyrus, in the medial anterior temporal lobes bilaterally (in proximity to the head of the hippocampus, fusiform gyrus and amygdala), posteriorly in a small cluster in close proximity to the tail of the hippocampus, and finally in bilateral premotor cortices and left SMA (Figure 41A).

5.5.4.2 Face Finder (<rest)

Performance of the Face Finder game was associated with activation in bilateral posterior inferior occipital gyri, angular gyri (AG) and right lingual gyrus ($p<0.05$ FWE corrected) (Figure 41B). At a lower threshold activation was also seen in bilateral middle occipital gyri, superior parietal lobe (SPL), SMA, precentral gyri, Pars Opercularis (below the surface of the cortex, on the border of BA 4) and the IFS (on the border of BA 44/45 and the middle frontal gyrus). Left hemisphere activation was seen in the putamen, anterior insula and opercular cortex, as well as posterior fusiform gyrus. There was also localised activation in the right midbrain.

5.5.4.3 Semantic and syntactic processing from orthographic input (RN>FF)

There was no significant activation in sentence-level semantic or syntactic processing regions when correcting for multiple comparisons ($p<0.05$, FWE corrected. At a lower threshold ($p<0.001$, $k>10$, uncorrected) significant activation was seen bilaterally in visual regions (calcarine and lingual gyri), as

well as the right posterior STG and left vOT, temporal pole, and Pars Triangularis and Orbitalis (BA 45/47) (Figure 41B).

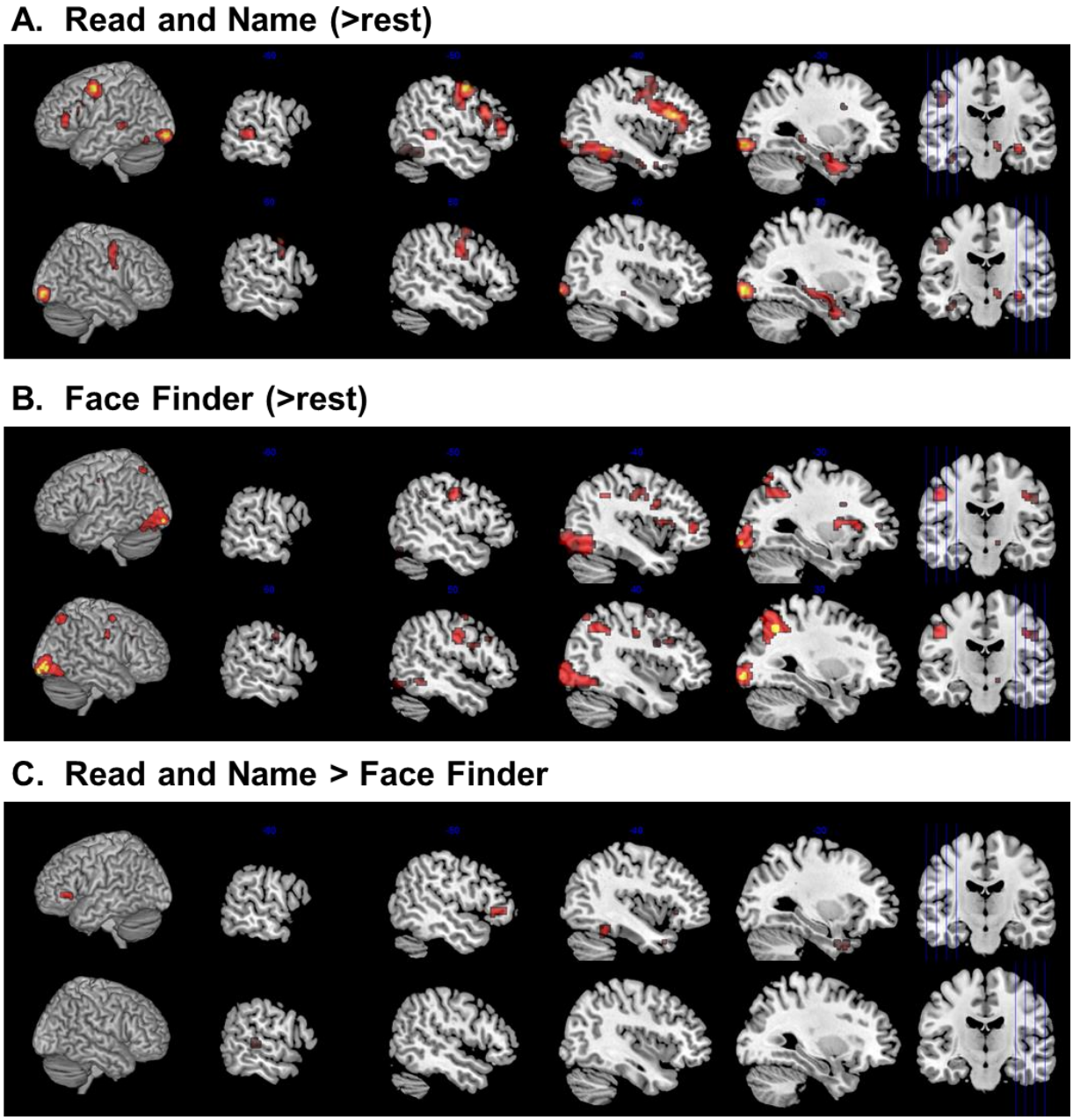


Figure 41: Group activation (N=18, 9-16 years) for reading comprehension and word retrieval (A), visual processing of symbol strings and word repetition (B) and for semantic and syntactic processing from orthography (C). For each contrast activation is shown for the left (top) and right (bottom) hemispheres, on a rendered image of the brain (left) and on five sagittal slices (right). Activation is shown at $p<0.05$ FWE corrected (yellow) and $p<0.001$ uncorrected (red).

5.6 Amodal semantic processing regions

Regardless of input modality, semantic processing during naming was associated with activation in the left vOT, left temporal pole, and right anterior fusiform gyrus (p<0.001, k>10 uncorrected). At a lower threshold (p<0.005, k>10 uncorrected) activation was also seen in left Pars Triangularis (BA 45), left anterior fusiform gyrus, and left entorhinal cortex (Figure 42). There was no significant activation at p<0.05 FWE corrected, even when using a small volume correction for expected significant differences in the temporal and frontal lobes.

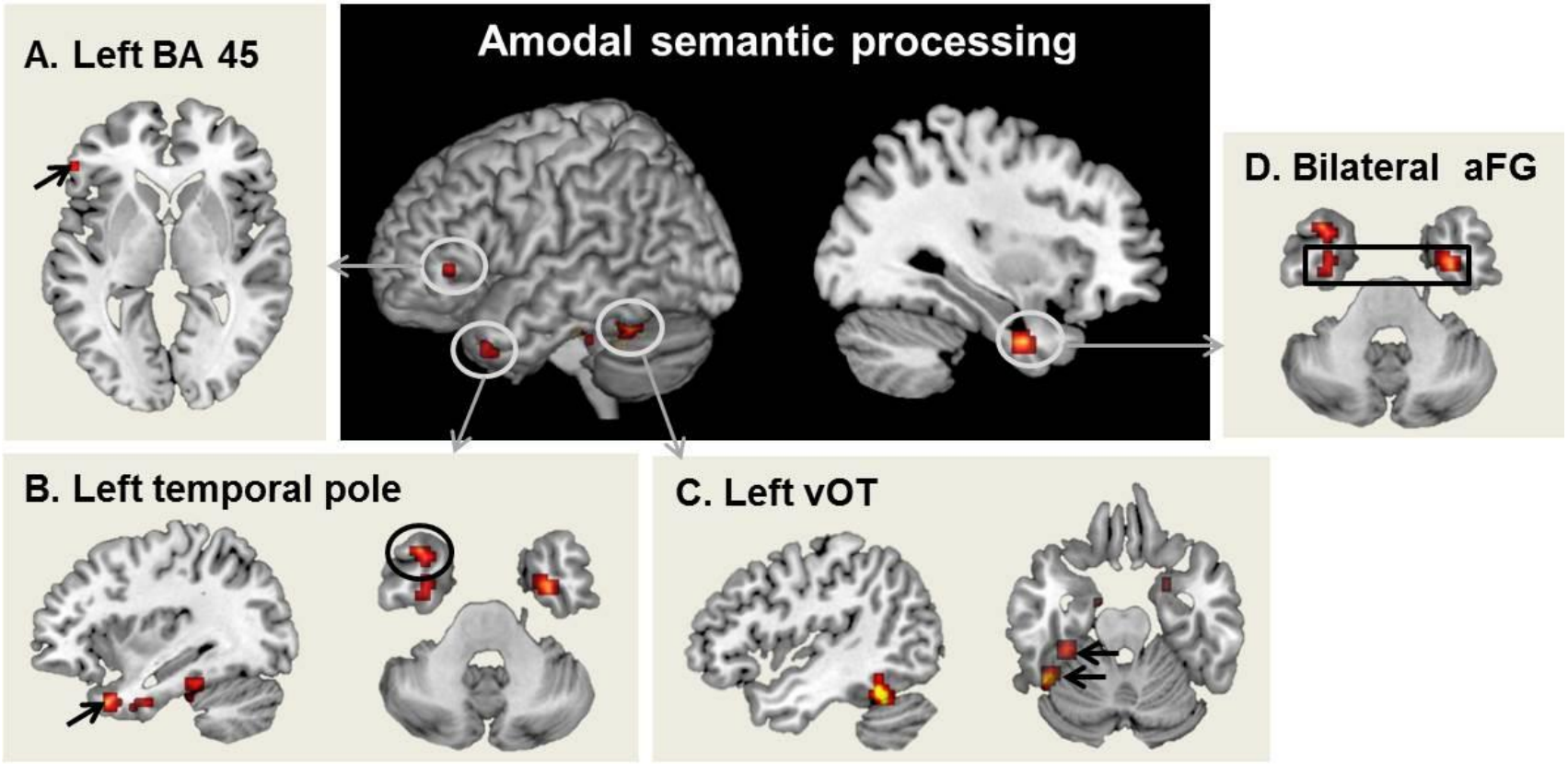


Figure 42: Amodal semantic processing network. Activation common to semantic processing during naming from pictorial, phonological and orthographic input is shown on a rendered image of the left hemisphere and sagittal slice of the right hemisphere (black box). A) Left inferior frontal gyrus peak in Pars Triangularis (BA 45), B) Left temporal pole peak on a sagittal slice (arrow) and axial slice (circle), C) Two peaks in the left ventral occipito-temporal gyrus (vOT) shown on an axial slice (arrows) with the lateral cluster also displayed on a sagittal slice, D) Bilateral anterior fusiform gyrus (aFG) peaks. Activation shown at $p<0.001$ (yellow) and $p<0.005$ (red) uncorrected.

6. DISCUSSION

6.1 Scan failures

Relative to previous studies of functional MRI in children and adolescents, failure rates for the Panda Games were relatively low, ranging from 16% (for PN) to 42% (for PD), compared to 31-57% in previous studies (Byars et al., 2002; Yerys et al., 2009). The majority of scan failures were caused by in-scanner movement (and associated artefact), accounting for 100%, 61%, 82% and 50% of scan failures for PN, PD, LN and RN, respectively. The contribution of movement to scan failure was higher in this study relative to others (~16% in Yerys et al., 2009 and Croft et al., 2013). Movement may have accounted for a greater proportion of scan failures due to their being relatively few failures caused by poor performance (0% of PN failures, 8% of LN failures, 4% of RN failures and 25% of PD failures). As expected, in-scanner movement was higher in younger children; a vulnerable age-group for scan failure (Byars et al., 2002). Younger children also showed more large sporadic head movements (e.g. head jerks). Poorer experience inside the scanner and headphone discomfort during scanning were also associated with increased in-scanner movement. *In-scanner experience* was measured according to how much participants enjoyed the Panda Games and how difficult they found them, how claustrophobic and how comfortable they felt during scanning, as well as how long they felt they were inside the scanner. It may be possible to reduce scan failures in children and adolescents by improving the physical comfort of participants during scanning (including the use of child-friendly headphones), as well as making fMRI tasks engaging and keeping scan times short. In this sample of healthy children, movement was not associated with executive functioning (including measures of behavioural regulation) or overall ability (FSIQ). However this may not be the case in children with epilepsy (Croft et al., 2013).

Mock scanning did not significantly reduce in-scanner movement or scan failure rates in this sample. The effect of mock scanning on these variables may have been reduced due to the very long scanning times, during which several patients felt physical discomfort (as reported on the in-scanner questionnaire measure).

However, children who underwent mock scanning did show improved in-scanner performance across the battery of tasks, despite spending an equal time practicing the tasks compared to the no mock group. Further, the general response from participants and families was that the pre-scan preparation work was valued.

6.2 Typical networks supporting performance of the Panda Games

Patterns of activation associated with performance of each of the Panda Games are described below. Functional interpretations are provided in brackets for each region and, unless otherwise stated, are based on consistent structure-function mappings reported in the mature language system (Price 2012) and described in the functional neuroanatomical model of language (introduced in section 3.2 and shown in Figure 30 and Figure 31). My main findings regarding localisation of core linguistic (semantic and syntactic) processing are summarised for each task in Figure 43-Figure 46, where my findings are highlighted (orange) in the context of predictions made by the functional neuroanatomical model (white).

6.2.1 Picture Naming

Group map activation for Picture Naming (compared to rest) included; bilateral occipital cortex (early visual processing), bilateral temporal poles (semantic processing), left premotor cortex (action planning), SMA (motor execution) and bilateral precentral gyri (orofacial motor activity). Activation was also seen bilaterally along the length of the fusiform gyrus, from posterior occipital regions to anterior basal temporal regions. The role of the fusiform gyrus in language is discussed further in section 5.3.1.

6.2.1.1 Semantic processing during Picture Naming

By contrasting PN and CNw, I was able to localise semantic processing regions in line with predictions of the functional neuroanatomical model (Figure 43). I also found additional activation of bilateral occipital cortex, which may reflect increased visual processing in the PN condition versus the CNw condition.

Previous studies of word generation report activation in lateral temporal regions (Price, 2012), which is less lateralised to the left hemisphere in children (Everts et al., 2009; Holland et al., 2001; Karunanayaka et al., 2010; Szaflarski, Holland, et al., 2006; Szaflarski, Schmithorst, et al., 2006). My findings provide new evidence to suggest semantic aspects of word generation are left lateralised in children (even those as young as 5 years of age), and that these core linguistic aspects of task performance involve the posterior ITG and temporal pole specifically. My findings also provide novel evidence in children, that the vOT is increasingly engaged in naming pictures than naming colours; as has been shown in adults (Price & Devlin, 2003). While adults show activation specific to the left vOT however (Cai, Lavidor, Brysbaert, Paulignan, & Nazir, 2008; Seghier & Price, 2011, 2013), my findings suggest children recruit bilateral vOT to support picture naming.

To summarise, my findings suggest semantic aspects of picture naming are supported by left lateralised lateral frontal and temporal cortex, similar to adults, and also bilateral vOT, which are left lateralised in adults. Age-related changes in the vOT require further investigation.

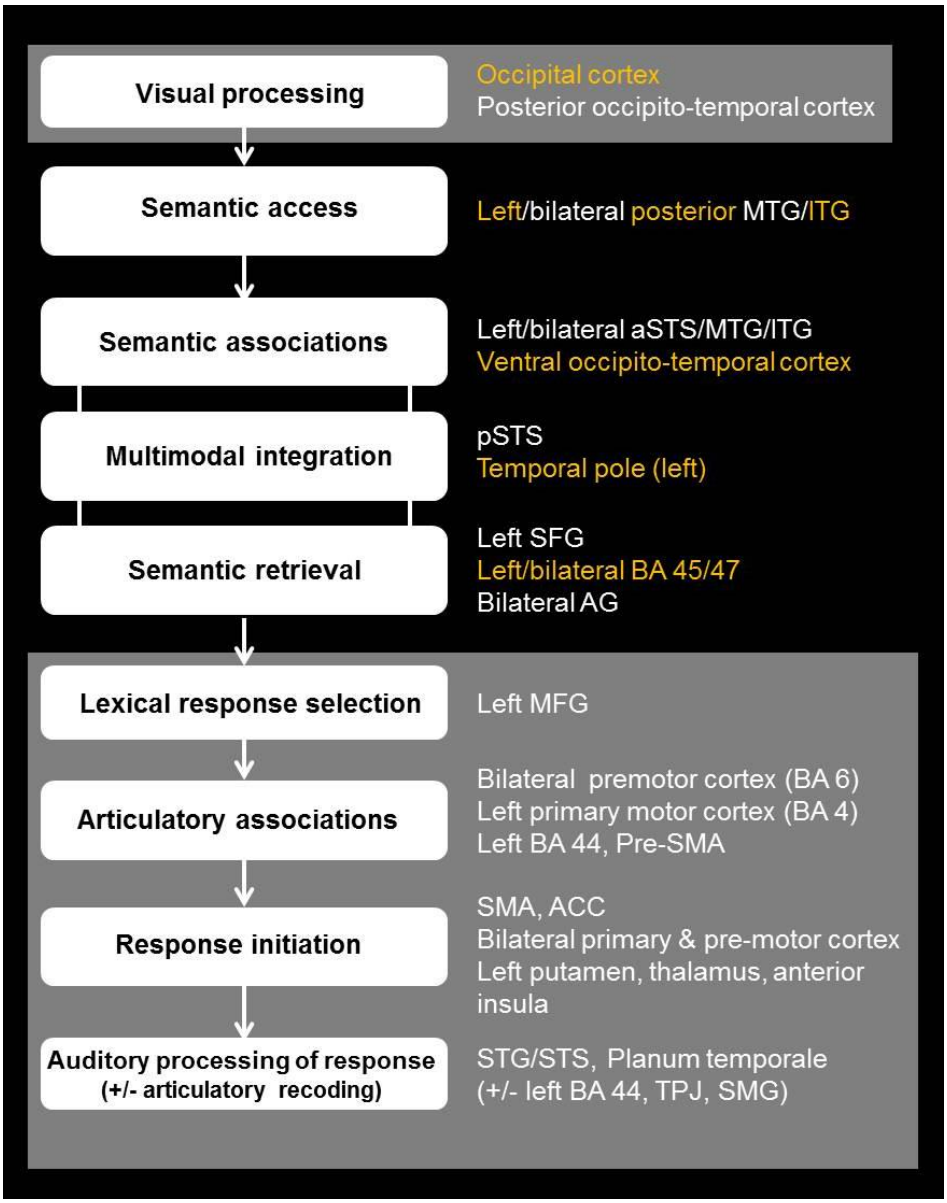


Figure 43: Hypothesised and observed activation associated with semantic processing during the Picture Naming task. Hypothesised regions of activation are listed (white text) for all processing stages specified in the functional neuroanatomical model of language proposed by Price 2012 and involved in Picture Naming (white boxes). Processes and brain regions hypothesised to be involved in the Picture Naming task (black box) and the Colour Naming with words baseline task (grey box) are shown. Processes and brain regions hypothesised to be involved in semantic aspects of picture naming are specific to the black box, and observed activations resulting from the contrast PN>CNw are highlighted in orange for these processes.

6.2.2 Picture Describing

Findings from the Picture Describing task are shown in the context of predictions made by the amended *Blue print of the speaker* model presented in Figure 33, and literature reviewed in section 3.2.2 (Figure 44). As hypothesised, strong activation was found associated with visual processing (bilateral dorsal occipital cortex) and articulation (SMA and precentral gyri) during performance of the Picture Describing task (compared to rest). Extensive activation was also seen in a bilateral ventral temporal system, which extended from the posterior ventral occipito-temporal cortex along the fusiform gyrus to the temporal pole, and included the hippocampus. Neither the hippocampus nor anterior temporal regions were predicted by Levelt's model (Figure 33). However, the temporal pole has been implicated in semantic processing (Binder et al., 2009), and may provide an interface between semantic representations and cognitive control systems (Binney, Parker, & Ralph, 2012). How the hippocampus contributes to sentence-level expressive language requires further investigation.

6.2.2.1 Semantic processing during Picture Describing

Within the ventral temporal network discussed above, activation in the vOT was specifically associated with semantic aspects of sentence generation (Figure 44). Activation in this region (but only in the left hemisphere) was the most consistently reported area of activation in adults (section 3.2.2). Activation of Broca's area was variable in my sample, and was only visible at a very low statistical threshold, also similar to adults. It is possible activation in this region is modulated by syntactic and semantic demands (Brauer & Friederici, 2007). There may also be multiple routes to sentence generation, involving different substrates; similar to the reading network (Richardson, Seghier, Leff, Thomas, & Price, 2011). My results suggest children consistently recruit a bilateral ventral temporal network for generating simple sentences based on visual input. This network includes regions involved in picture naming, and may provide a relatively fast and automatic strategy for generating sentences when semantic and syntactic demands are low, and generation is based on external visual input.

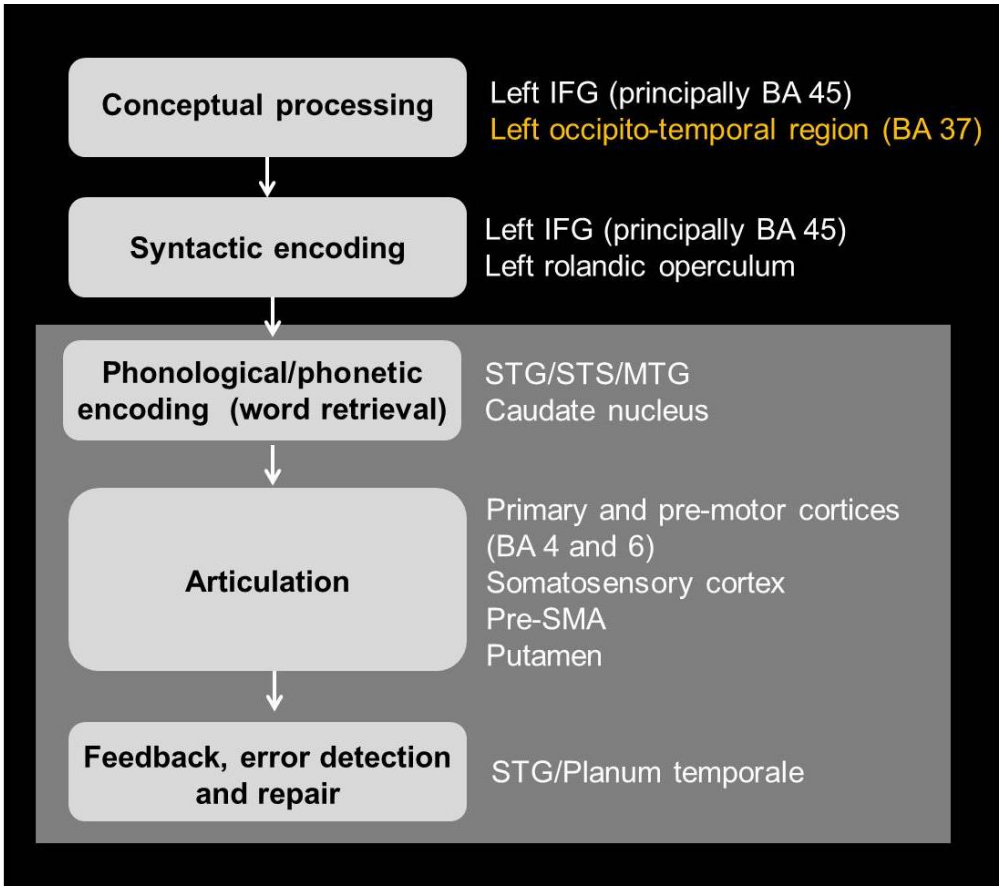


Figure 44: Hypothesised and observed activation associated with semantic processing during the Picture Describing task. Hypothesised regions of activation are listed (white text) for all processing stages specified in Levelt's 'blue print of the speaker' model and involved in Picture Describing (light grey boxes). Processes and brain regions hypothesised to be involved in the Picture Describing task (black box) and the Colour Naming with sentences baseline task (grey box) are shown. Processes and brain regions hypothesised to be involved in semantic aspects of picture describing are specific to the black box, and observed activations resulting from the contrast PD>CNs are highlighted in orange for these processes.

6.2.3 Listen and Name

Children and adolescents recruit a broad network of regions to support performance of the Listen and Name Game, including auditory comprehension regions previously reported children (Ahmad et al., 2003; Berl et al., 2010; Lidzba et al., 2011; Schmithorst et al., 2006; M. Wilke et al., 2005). In adults, these

regions have consistently been associated with auditory processing (bilateral STG), semantic processing (bilateral STG, temporal pole and ventral occipito-temporal cortex), semantic retrieval (left Pars Triangularis and Orbitalis), auditory imagery (left planum temporale), speech production (bilateral thalamus and left anterior insula), motor execution (SMA) and orofacial motor activity (bilateral precentral gyri). I also found activation beyond the functional neuroanatomical model, including bilateral hippocampi and the medial superior frontal gyrus. These regions are involved in memory (Eichenbaum, 2000) and cognitive control (Rushworth, Walton, Kennerley, & Bannerman, 2004), and have been reported during auditory comprehension in children previously (Balsamo et al., 2006; Schmithorst et al., 2006).

6.2.3.1 Semantic processing during auditory comprehension

Similar to PN and PD, semantic and syntactic processing during auditory comprehension was associated with ventral temporal activation which extended from posterior lingual gyri and vOT, along the fusiform gyrus to the temporal pole. Activation in vOT and the temporal pole was bilateral, however activation between these regions (the anterior fusiform gyrus; discussed later in section 6.3.2) was only found in the left hemisphere. Activation was also found in Broca's area (left BA 45 and 47), bilateral anterior STG/STS and left thalamus, as predicted by the functional neuroanatomical model (Figure 45). Regions not predicted by the model included right hippocampus and medial superior frontal gyrus. Increased activation of these regions during LN compared to AG suggests they contribute specifically to the processing of intelligible speech.

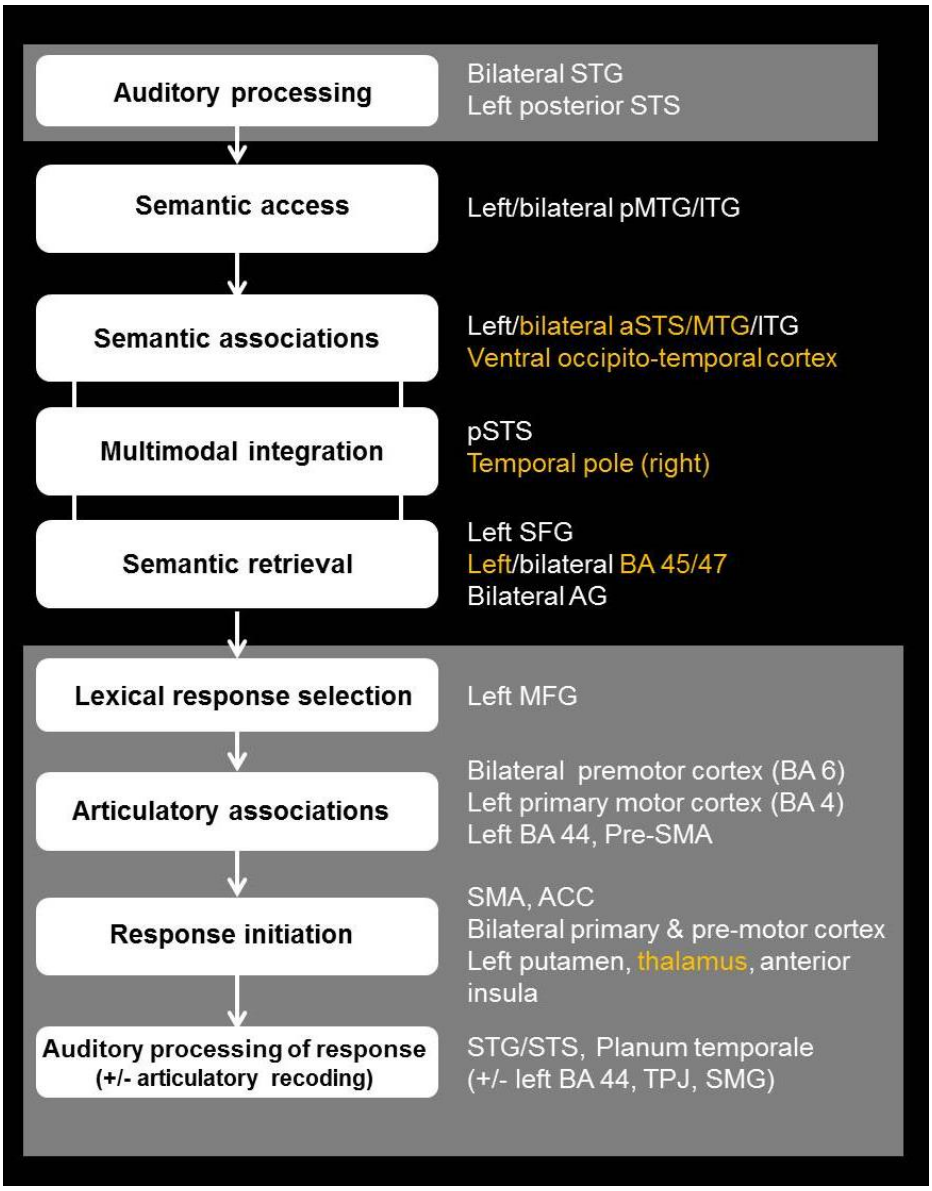


Figure 45: Hypothesised and observed activation associated with semantic processing during the Listen and Name task. Hypothesised regions of activation are listed (white text) for all processing stages specified in the functional neuroanatomical model of language proposed by Price 2012 and involved in word retrieval from phonological input (white boxes). Processes and brain regions hypothesised to be involved in the Listen and Name task (black box) and the Alien Game baseline task (grey box) are shown. Processes and brain regions hypothesised to be involved in semantic aspects of auditory comprehension are specific to the black box, and observed activations resulting from the contrast LN>AG are highlighted in orange for these processes.

6.2.4 Read and Name

As hypothesised, during the Read and Name game activation was seen in regions previously associated with reading, including; bilateral occipital regions (visual processing), left posterior MTG/ITG (direct semantic access), left occipito-temporal cortex (semantic associations), bilateral temporal poles (multimodal semantic integration), left Pars Orbitalis (semantic retrieval), left Pars Opercularis and bilateral precentral gyri (articulatory/phonological processing), and the SMA (articulation). These findings show children as young as 8 years of age are able to obtain direct semantic access via the left posterior MTG/ITG and occipito-temporal cortex, similar to adults (Price, 2012) and contrary to previous findings (Turkeltaub et al., 2013). There was no evidence for the use of an immature reading strategy indicated by right occipito-temporal activation (Orton, 1935). The left occipito-temporal cluster included a peak of activation in the classic 'visual word form area' (Cohen et al., 2002). However, activation in this region may not be specific to reading (Price & Devlin, 2003). Indeed, activation in proximity to this region was seen across the Panda Games task battery (this is discussed further in section 6.3.1).

As well as identifying regions hypothesised by the functional neuroanatomical model, I also identified bilateral anterior and ventro-medial activations in the temporal lobes, including the anterior parahippocampal gryus/hippocampus; similar to regions reported for LN. These provide further evidence for a role of the medial temporal cortex in sentence-level language, which warrants further investigation.

6.2.4.1 Semantic processing during reading comprehension

By contrasting the reading task with a symbol search task, I was able to localise semantic and syntactic processing regions consistent with those found in adults (Figure 46). Unlike previous studies in children (Berl et al., 2010; Shaywitz et al., 2007; Turkeltaub et al., 2003), I found semantic processing during reading is predominantly associated with left hemisphere activation, including; Broca's area (BA 45/47), the temporal pole, and occipito-temporal

cortex. The posterior STS was only active in the right hemisphere however. The posterior STS is associated with multimodal integration in the functional neuroanatomical model, and the right pSTS has specifically been associated with development of pre-reading skills in children (Rapid Automatised Naming) (Turkeltaub et al., 2013). In addition to these regions, I found activation in bilateral calcarine sulcus, and the left medial and posterior lingual gyrus. The latter has been associated with global (as opposed to local) processing of visual forms (Fink et al., 1996), and may reflect whole-word processing during reading. Activation in this region therefore provides further evidence that children were able to use a whole-word strategy to gain direct access to semantic processing during performance of the RN game. The developmentally-appropriate task design in the Panda Games may have enabled identification of this more efficient direct route to reading, which may be obscured in young children when the same task is applied across a broad age range.

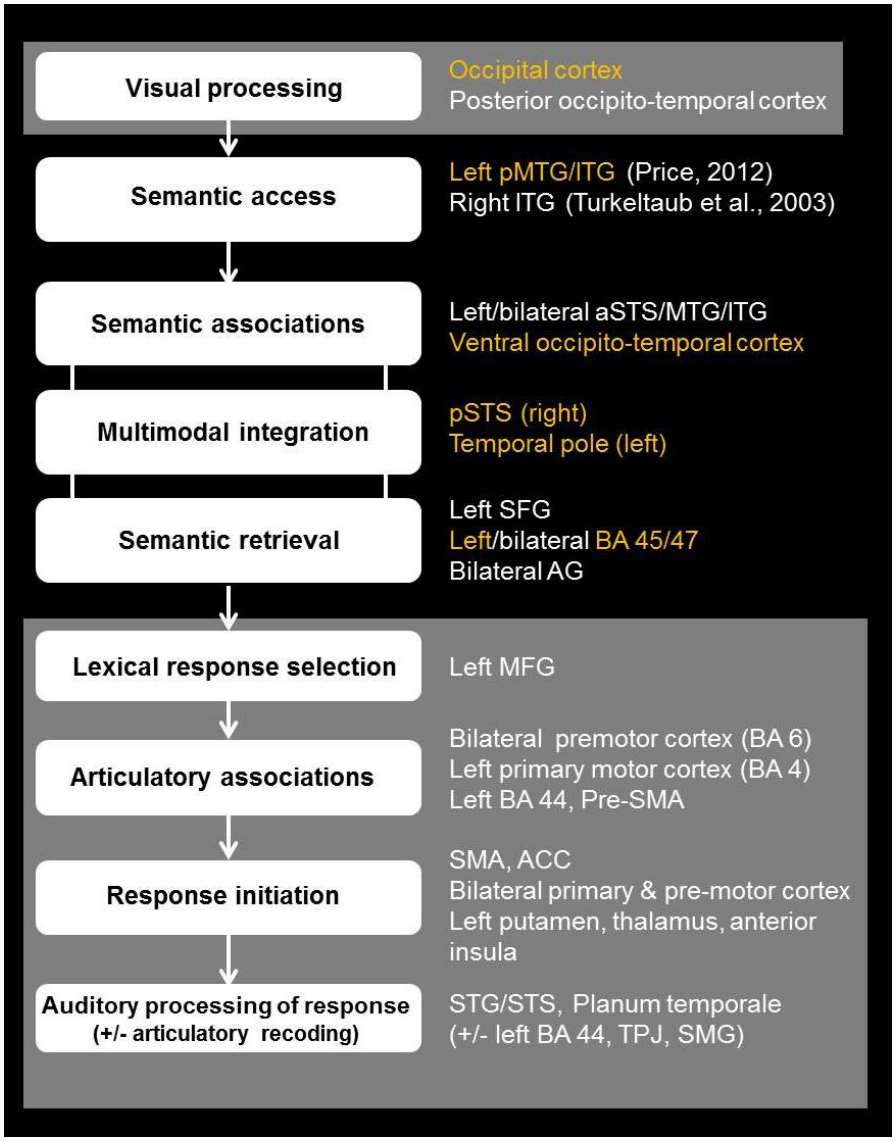


Figure 46: Hypothesised and observed activation associated with semantic processing during the Read and Name task. Hypothesised regions of activation are listed (white text) for all processing stages specified in the functional neuroanatomical model of language proposed by Price 2012 and involved in word retrieval from orthographic input (white boxes). Processes and brain regions hypothesised to be involved in the Read and Name task (black box) and the Face Finder baseline task (grey box) are shown. Processes and brain regions hypothesised to be involved in semantic aspects of auditory comprehension are specific to the black box, and observed activations resulting from the contrast RN>FF are highlighted in orange for these processes. NB: Activation in the left MTG/ITG was only significant for contrast RN>rest.

6.3 Amodal semantic processing regions in children

Having summarised the unique activation patterns associated with each of the Panda Games, I now discuss consistent findings across tasks. In doing so I highlight parts of the language system which are most reliably identified using the Panda Games battery, regardless of age. This information should help identify pre-surgical candidates for whom these tasks may be most helpful.

Across all tasks, I identified activation in a basal ventral temporal network, which extended from posterior vOT along the fusiform gyrus to the temporal pole and – for sentence-level tasks – included medial temporal regions. Activation in this network was bilateral for word/sentence generation from pictures (PN and PD) and was left lateralised for comprehension (LN and RN). This network also included more variable activation in semantic processing sub-regions of classic Broca's area (BA45 and BA 47). Regions in this network were also identified in my analysis of amodal semantic processing (Figure 42), and are discussed below. My findings are summarised in relation to a cognitive model of naming which is helpful in defining different semantic processes, and their impairment in patients with structural damage in specific parts of the language network (Figure 49), this account is discussed at length in De Leon et al,. (2007), from which I adapted the schematic.

6.3.1 Ventral occipito-temporal cortex

The left vOT was the most consistently active region within the language system during performance of the Panda Games, and was identified as an amodal semantic processing region. Activation in the right vOT was also seen during performance of the sensory-baseline tasks. The vOT refers to a sub-region within the occipito-temporal cortex (OT), which includes both the lateral posterior ITG and the mid-fusiform gyrus. I use the term vOT to specifically refer to the mid-fusiform part of the OT, which lies on the border of the occipital and temporal lobes (Figure 48A), and is also known as Brodmann's area 37 (Brodmann, 1909). The fusiform gyrus in its entirety actually runs along the base of the brain, from visual cortex to the temporal pole (Figure 47). I found both an anterior and a

posterior cluster in the left vOT involved in amodal semantic processing; a posterior cluster (*y*=-49) and anterior cluster (*y*=-34).

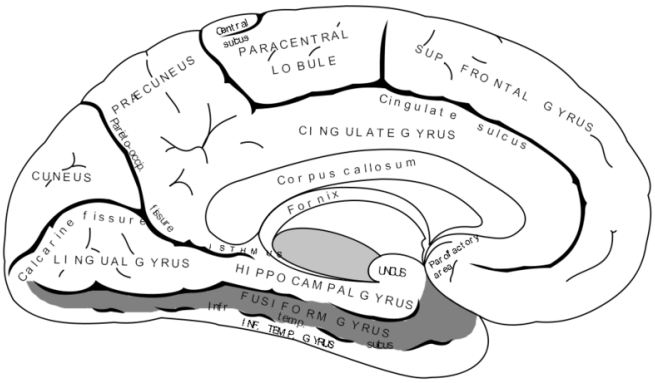


Figure 47: The fusiform gyrus.

Cohen and colleagues claim a posterior region in vOT (with Talaraich co-ordinates of ~y=-54 to -60) is specifically involved in reading as a *visual word form area* (VWFA), while the more anterior aspect of vOT (~y=-30 to -54) is involved in amodal, abstract representation (Cohen et al., 2002). However, previous fMRI findings (Price & Devlin, 2003) and my own findings presented here, suggest the posterior vOT responds across a range of tasks; some of which do not involve visual word form processing at all. These include picture naming, generating a verb to pictures, viewing pictures, colour naming, as well as auditory and tactile word processing tasks (Price & Devlin, 2003). I provide novel evidence to suggest this region is also involved in sentence generation and auditory comprehension (Figure 48B).

Intracranial mapping studies refer to the vOT (specifically the region located ~7-9cm from the tip of the temporal pole) as the basal temporal language area (BTLA) (Davies et al., 1994; Krauss et al., 1996; Luders et al., 1989; Luders et al., 1986; Luders et al., 1991). These studies provide evidence to suggest the vOT cortex is critical to both naming and comprehension, with approximately 40-100% of patients showing critical language sites in the left vOT (Davies et al., 1994; Krauss et al., 1996; Luders et al., 1991). Further, the prevalence of critical

language sites found during reading in this region is comparable to both Broca's and Wernicke's area (Schaffler, Luders, & Beck, 1996); both of which are considered core language regions in the classic neurological model of language.

Consistent activation in both the posterior and anterior aspects of the vOT across the Panda Games task battery in children, as well as previous findings from fMRI and ICM studies in adults, suggest that this region supports cognitive processing common to a range of language tasks. Over 15 years ago Buchel and colleagues suggested this region promotes activation in distributed regions of the cortex which – collectively – form a concept (semantic retrieval) or provide lexical access to phonological and orthographical word forms (Buchel, Price, & Friston, 1998). More recent reviews of the fMRI literature provides further evidence to my own, to suggest the left vOT forms part of an amodal semantic processing network (Binder et al., 2009). Specifically, patterns of naming errors in adults with stroke suggest the left vOT is involved in amodal lexical access (DeLeon et al., 2007). The specific role of the left vOT in naming is summarised in Figure 49.

I provide evidence to suggest the left vOT should be considered a core language region, involved in amodal semantic processing from as young as 5 years of age. Further, I present evidence from adults to suggest this region is critical to language. I therefore hypothesise the use of the Panda Games to accurately identify activation in this region will aid in the pre-surgical decision making process when the surgical target implicates surrounding cortex (see Patient F presented in Chapter 6 for an example). Surprisingly, the contribution of this region to language development has received little investigation. I address this directly in Chapter 8.

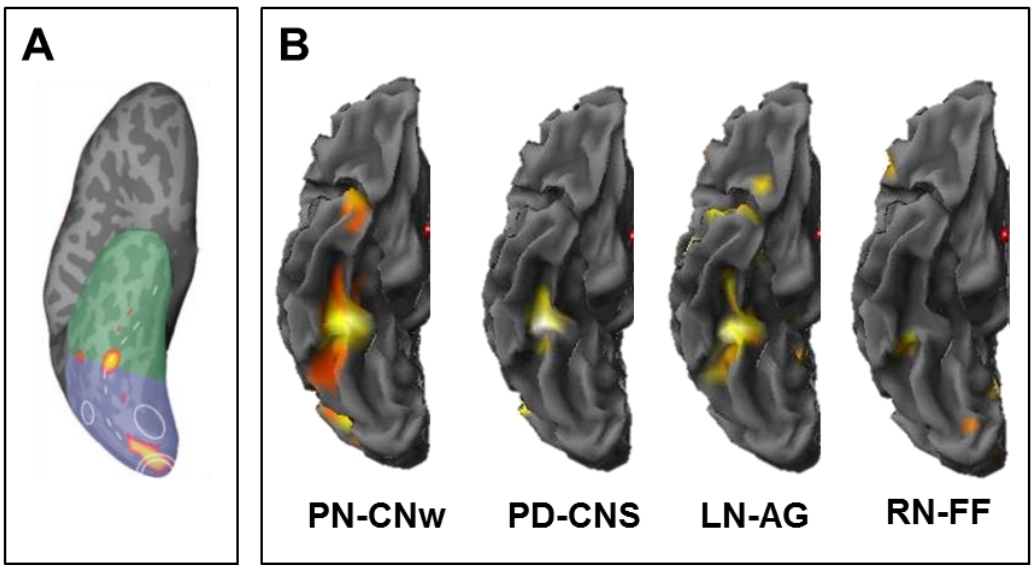


Figure 48: Location of ventral occipito-temporal cortex (vOT) activation in the left hemisphere which has previously been labelled the 'Visual Word Form area' (A), and is activated across tasks in the Panda Games battery (B). Figure A is taken from Price and Devlin, (2011). Figure B shows activation associated with semantic processing during performance of Picture Naming (PN>CNw), Picture Describing (PD>CNs), Listen and Name (LN>AG), and Read and Name (RN>FF).

6.3.2 Temporal pole and anterior fusiform gyrus

I found two anterior temporal regions involved in amodal semantic processing in children (Figure 42). Most of the evidence for involvement of the anterior temporal lobes in semantic processing comes from patients with semantic dementia; defined by progressive atrophy in bilateral anterior temporal regions (Hodges & Patterson, 2007; Hodges et al., 1992; Jefferies et al., 2008; Mummery et al., 2000). These patients demonstrate semantic errors which indicate progressive loss of semantic memory (the concepts of words), suggesting this region acts as a store for semantic representations (Figure 49). Functional MRI studies in healthy adults also suggest these regions provide an amodal hub in the semantic processing network (Binder et al., 2009). I provide new evidence to suggest that the anterior temporal lobes are involved in the semantic processing network as early as 5 years of age, specifically the left temporal pole and bilateral anterior fusiform gyri. In patients with semantic dementia, semantic abilities are

predicted by the degree of atrophy in the anterior fusiform gyrus specifically (Mion et al., 2010). The left and right anterior fusiform regions also show evidence for a degree of material specificity in these patients. This was not evident in my developmental cohort, and raises the possibility of increasing specialisation within this system with age.

The anterior temporal lobes are often resected in children with drug-resistant temporal lobe epilepsy. In children, this procedure has been associated with an initial decrease in verbal memory ability at short-term follow-up (Gleissner, Sassen, Schramm, Elger, & Helmstaedter, 2005; Szabo et al., 1998), although long term cognitive outcome is good (Skirrow et al., 2011). Consistent activation of the anterior temporal lobes in the Panda Games provides encouraging evidence to suggest these tasks will provide information to aid the paediatric neurosurgical imaging process in cases where the surgical target lies within anterior temporal regions.

6.3.3 Left Pars Triangularis and Pars Orbitalis

Regions within Broca's area which are associated with semantic retrieval (left Pars Triangularis and Orbitalis) were active in all of the Panda Games, although activation was variable for sentence generation, similar to adults (Blank et al., 2002; Muller et al., 1997). In contrast I found activation of the Pars Opercularis was limited to the reading comprehension task (RN), consistent with predictions of the functional neuroanatomical model: that this region supports phonological/articulatory processing (Price, 2010). Frontal regions support goal-orientated semantic retrieval of information necessary for lexical-semantic access, as shown in Figure 49. Frontal activation is thought to be variable and more bilateral in young children (Everts et al., 2009; Holland et al., 2001; Karunanayaka et al., 2010; Szaflarski, Holland, et al., 2006; Szaflarski, Schmithorst, et al., 2006; Vannest et al., 2010). However, using a high-level contrast to localise semantic processing, I provide evidence to suggest semantic retrieval produces consistent and left lateralised activation in Pars Triangularis and Orbitalis in children, as in adults.

6.3.4 Medial temporal cortex

A small cluster of activation in proximity to the entorhinal cortex was identified as part of the amodal semantic processing network in children. Additionally, more extensive activation in parahippocampal and hippocampal regions was observed during sentence comprehension (LN and RN) and generation (PD).

Medial temporal regions including the entorhinal cortex, parahippocampal gyrus and hippocampus together form the medial temporal memory system. Previous studies have shown medial temporal activation during comprehension in children (Balsamo et al., 2006; Schmithorst et al., 2006) and there is evidence for medial temporal involvement for learning the rules of grammar (Opitz & Friederici, 2004). Further, evidence from children with epilepsy emphasises the role of the medial temporal cortex in verbal memory (Gabrieli, Cohen, & Corkin, 1988; Szabo et al., 1998) and also language reorganisation (Liegeois et al., 2004). The contribution of the hippocampus to the 'online' use of language in adults is also becomingly increasingly acknowledged, based on evidence from patients with amnesia (Kurczek, Brown-Schmidt, & Duff, 2013). Indeed, the memory system has been specified as one of the foundations for naturalistic language in some models; e.g. the Memory, Unification and Control model (Hagoort, 2013).

My findings provide novel evidence to suggest children recruit additional resources from the memory system to support semantic processing. The interplay of language and memory systems may be an important aspect of language development, although this has not yet been investigated. Crucially, understanding the effect of cross-domain system interactions on the localisation and lateralisation of language could provide information of value in the pre-surgical setting. Further, it may shed new light on the functional neuroanatomical correlates of language impairment common in children with epilepsy (Parkinson, 2002).

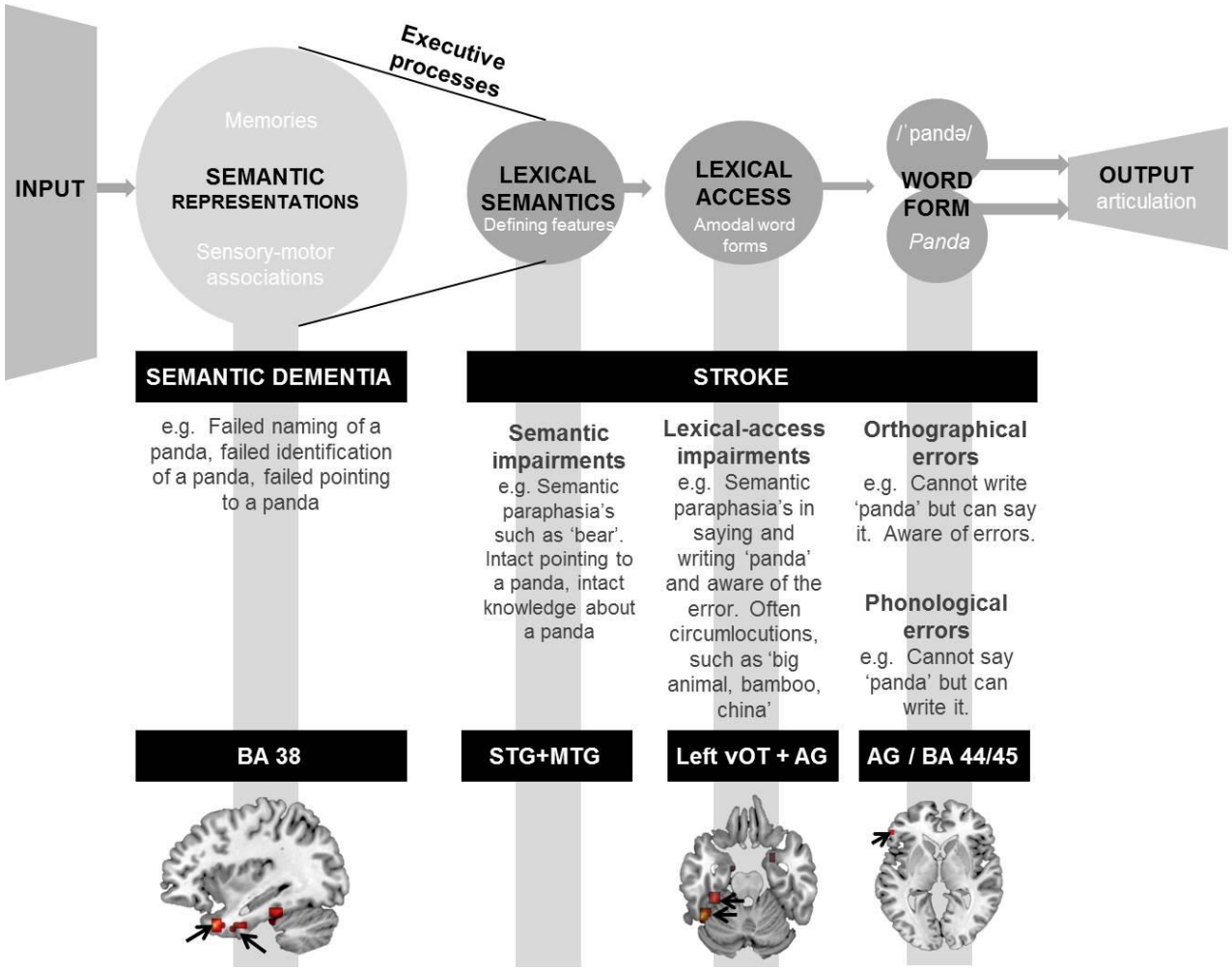


Figure 49: The semantic processing stream supporting naming of a panda, and the neural correlates associated with each stage based on findings from patients (black boxes) and from this chapter (brain images). The cognitive architecture supporting naming is shown (top), based on data reviewed and presented in De Leon et al., 2007. Types of semantic processing impairments are described (grey) and their neural correlates are listed (black boxes) for patients with semantic dementia (based on Hodges et al., 1992 and Mion et al., 2010) and for patients with stroke (based on De Leon et al., 2007). My own findings in healthy children are shown at the bottom of the schematic, and show close correspondence with critical regions reported in patients.

7. CONCLUSIONS

The Panda Games successfully identified ventral systems engaged by receptive and expressive language in children aged 5-16 years. These networks included regions hypothesised by adult models. Principally, the Panda Games identified a ventral temporal network which extended from vOT to the temporal pole. Regions within this network supported amodal semantic processing, and have been shown to be critical to language function in adults. I therefore provide evidence for a set of core language regions in children, which include the left temporal pole, vOT, Pars Triangularis and bilateral anterior fusiform gyri. These regions require further investigation in children with epilepsy in terms of predicting language impairment and post-operative language outcome. The medial temporal lobe memory system also contributes to this semantic processing network in children. The interaction between memory and language systems highlights a new avenue for future research which could shed light on the development of the language system, the potential for reorganisation of function, and language impairment.

Chapter 6: Pilot of the Panda Games in a representative sample of epilepsy surgery candidates aged 5-16 years.

1. ABSTRACT

Children with epilepsy who are considered for neurosurgery often have cognitive impairment and behavioural comorbidities, making fMRI assessments of language particularly challenging in this population. In this chapter I aimed to pilot the new Panda Games fMRI task battery in a representative sample of children with drug resistant epilepsy who are being considered for neurosurgical intervention. Thirteen children with drug resistant epilepsy (aged 5-16 years) performed the Panda Games during fMRI. Sensitivity and specificity of results were compared between the new Panda Games battery and the current gold standard language task for presurgical fMRI (verb generation). Two patient cases were investigated in detail, to examine the clinical utility of the Panda Games, and their sensitivity and specificity in relation to invasive measures. When comparing word-level expressive language tasks, patients performed the Panda Games more accurately than verb generation. Patients performed sentence-level tasks in the Panda Games battery more poorly, with accuracy ranging from 8-83%. The Panda Games provided equivalent sensitivity and specificity to detect activation in language regions compared to verb generation. Case study 1 showed high sensitivity (0.75) and specificity (0.70) of the Panda Games in comparison to intracranial mapping. Case study 2 showed the Panda Games provide unique information regarding language localisation and lateralisation – not available from verb generation – which can contribute to the neurosurgical decision-making process. I conclude developmentally appropriate task design enhances performance accuracy in children with drug resistant epilepsy. Sentence-level tasks are more challenging for this population and must be slow-paced to enable accurate performance. The Panda Games produce sensitive and specific results regarding the localisation and lateralisation of language in comparison to both invasive and non-invasive current gold standard methods.

2. AIMS

In this chapter I investigate the feasibility of using the Panda Games in the pre-surgical setting, in terms of obtaining scans of useable quality (measured in terms of in-scanner task performance and movement) and attaining useful and reliable information regarding language localisation (measured in terms of sensitivity and specificity). I also present two case studies. The first provides an example of validation of the Panda Games in relation to the gold standard for mapping eloquent cortex (intracranial mapping) and demonstrates the issue of statistical thresholding and the type I and II error rates in fMRI. The second case study demonstrates the value of having a battery of tasks available such that fMRI can be individualised to the patient's specific clinical needs.

3. INTRODUCTION

I have shown the Panda Games are successful in identifying receptive and expressive language networks in healthy children aged 5-16 years, who show good performance on the tasks regardless of age. However, children with drug-resistant epilepsy are at risk of cognitive impairment (Bailet & Turk, 2000) and language impairment specifically (Parkinson, 2002). This population may find the Panda Games more challenging. I therefore conducted a pilot study to investigate the feasibility of using the Panda Games task battery in the clinical setting, in a representative sample of children with drug resistant epilepsy being considered for neurosurgical intervention. In order for the Panda Games to be useful in the pre-surgical setting it is important to know that 1) patients are able to perform the tasks well inside the scanner, 2) the Panda Games produce activation at the individual-level localised to target (language) regions, and 3) the task battery provides new information not available from current neuroimaging tasks. Another aim of the Panda Games was to enable the use of language fMRI with younger patients (< 7 years of age), by using a child-friendly fMRI protocol. Patients aged 5-6 years are therefore included in this pilot study. I also compared the results for the Panda Games with results from the current gold standard in pre-surgical fMRI (verb generation). This task has been well validated in relation to the WADA (Arora et al., 2009) and provides a good point of comparison. I hypothesise that performance

accuracy will be higher, and omission rates lower, for the Panda Games relative to VG, due to age-appropriate task design. I also hypothesise increased sensitivity to detect activation in basal temporal regions and the temporal pole for the Panda Games, based on normative data from Chapter 5. Finally, I hypothesise the Panda Games will provide new information regarding localisation and lateralisation of language in regions beyond those targeted by language fMRI tasks such as verb generation classic.

4. METHODS

4.1 Participants

Study information was sent to all patients referred for language fMRI as part of the pre-surgical decision making process at Great Ormond Street Hospital for Children NHS Foundation Trust (GOSH) from February 2013 - January 2014. Of the 32 patients referred for language fMRI during this period, six were not eligible to take part due to age (>16 years; N=3) or due to time restraints during the appointment (N=3). Of the 26 patients who met inclusion criteria, 14 (54%) applied to take part. Scanning was aborted during the standard clinical imaging protocol for one participant (aged 5 years) due to non-compliance and challenging behaviour; this participant is excluded from further analyses. In order to obtain a representative sample of children being considered for neurosurgery, all participants were included regardless of any specific clinical factors. The final sample included 13 participants with drug-resistant epilepsy (aged 6-16 years) (Table 20). The majority of patients (~70%) showed structural or EEG abnormality in proximity to perisylvian language cortex (Figure 50). One participant experienced continuous slow-wave spikes which were recorded throughout fMRI scanning using EEG-fMRI (Vulliemoz et al., 2009), and two experienced a seizure within 24 hours prior to scanning. The most recent seizure for the remaining participants occurred within the last week (N=4, 30%), the last month (N=2, 15%) or more than a month ago (N=2, 15%), or was not recorded (N=2). Neuropsychological data were available for 11 patients.

	Range	Mean	SD
Participant characteristics			
Age (years)	6 - 16	**11**	3.31
Sex (male:female)	-	**05:08**	-
Handedness (typical:atypical)	-	**10:02**	-
Neurosychological performance*			
Verbal IQ	57 - 112	**84**	17
Performance IQ	51 - 117	**91**	19
Epilepsy characteristics			
Age of onset (*years*)	1 - 11	4	3.26
Duration (*years*)	2 - 13	6	3.57
No. AEDs (*median*)	0 - 5	3	-
Seizure frequency			
Daily		6	**43%**
Weekly		4	**29%**
Monthly		3	**21%**
		N	**%**
Structural abnormality			**71%**
Left		7	**70%**
Right		1	**10%**
Bilateral		2	**20%**
EEG abnormality*			**100%**
Left		8	**57%**
Right		2	**14%**
Bilateral		1	**7%**

Table 20: Patient sample characteristics. The range, mean and standard deviation (SD) are shown for scalar data. The number (N) and proportion of participants (%) are shown for categorical data. Neuropsychological assessment was performed using the Wechsler Intelligence Scale for Children (fourth edition). Lateralisation of structural and EEG abnormalities are shown. Localisation of structural and EEG abnormalities are shown in Figure 50.

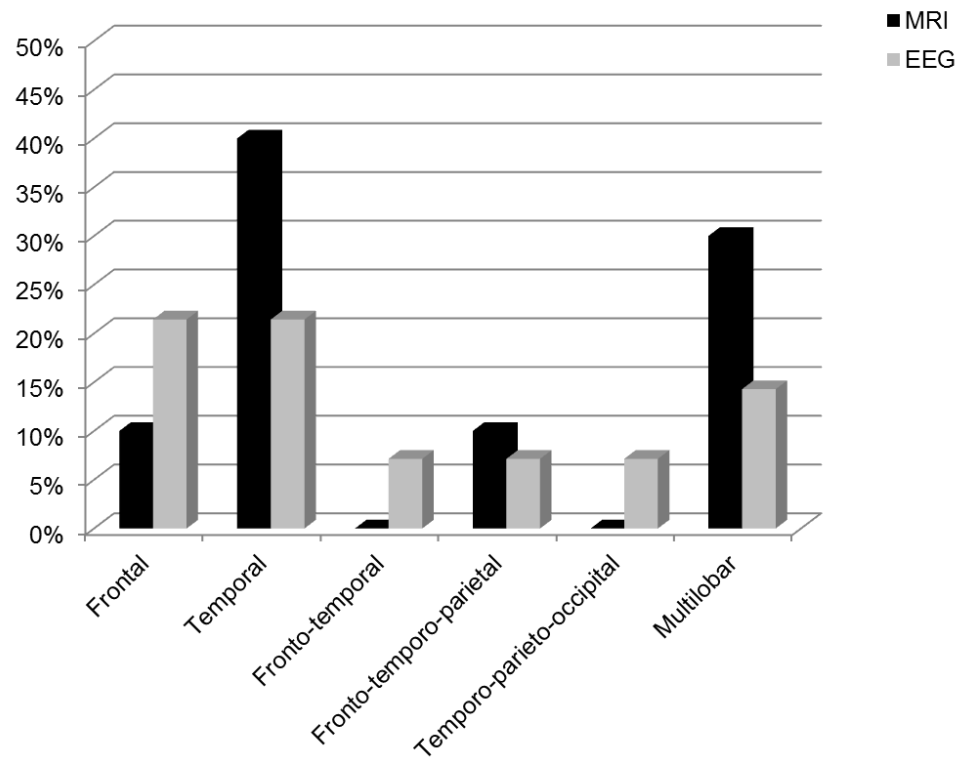


Figure 50: Location of structural abnormalities on MRI (black) and ictal activity on EEG (grey) in children aged 6-16 years with drug-resistant epilepsy. Structural and electrographical abnormalities involve the frontal or temporal lobes in 69% and 71% of participants, respectively.

4.2 Functional MRI

4.2.1 Panda Games

Structural and functional MR images were acquired as per Chapter 3. All participants performed the Panda Games (Chapter 3) after the standard clinical pre-surgical imaging protocol, which included: 1) overt verb generation, 2) covert verb generation, and 3) passive listening to stories. Data from the clinical protocol were acquired first and in the same order for every patient, due to clinical priority. The overt verb generation task is also reported in this chapter, to allow comparison of results from the Panda Games and results from a task currently used for pre-surgical assessments of language. Due to length of the scanning session four patients were not able to perform the Picture Describing task (PD), and one participant did not comply during the Listen and Name game (LN). In

total, 13 participants performed Picture Naming (PN), 9 performed PD, and 12 performed LN. For participants with performance accuracy >50% correct, only correct response trials were modelled. For participants with performance <50% correct on the language task or baseline, all response trials were modelled. The number of participants with performance <50% for each task was 0 (PN), 1 (PD), and 5 (LN).

4.2.2 Verb Generation

A well-validated block-design overt verb generation (to noun) paradigm (VG) was performed by all participants as part of the standard clinical pre-surgical imaging protocol (Croft et al., 2013). This task is used at GOSH to identify the language dominant hemisphere in surgical candidates. Children practiced the task prior to scanning and task information was included in the pre-scan information booklet designed for the Panda Games (Appendix 8.). During scanning auditory stimuli were transmitted through MRI compatible headphones. There were 8 VG blocks and 8 baseline blocks, each lasting 17.72 seconds. The total duration of the VG task was 4 minutes 41 seconds. There were a total of 32 stimuli, with four in each block (inter-stimulus interval = 4.32 seconds). During the VG blocks participants heard common nouns and were instructed to say a single verb associated with each noun ("a doing word that goes with it"). During the baseline condition participants listened passively to amplitude-modulated noise bursts of similar duration to the word stimuli in the active condition. One participant (aged 6 years) could not perform the VG task because it was too difficult and the white noise stimuli caused distress. In total 12 participants performed the VG task.

4.2.2.1 Pre-processing of verb generation

Only the first four blocks of verb generation were modelled at the first-level, to make the design matrix comparable to the Panda Games. Further, stimuli were modelled using an event-related design with an event duration of 0 seconds (comparable to PN). Temporal and dispersion derivatives were included in the model, as for the Panda Games, and all data were pre-processed using the motion fingerprint method (Wilke, 2012). For all participants with >50% correct (85%), only correct response trials were analysed. One participant

showed very poor performance (6% correct) due to tiredness and noncompliance. This participant generated semantically related nouns instead of verbs on 35% of trials, and repeated the stimulus or failed to respond on the remaining 59% of trials. For this patient all trials were modelled. A second patient showed poor performance (19% correct) because he generated word classes other than verbs on the large majority of trials. For this participant trials on which semantically related nouns, adjectives, pronouns and determiners (e.g. 'the') were generated were also deemed correct. This brought the total number of modelled trials up to 75% for this patient (only missed responses were excluded from the model).

4.3 In-scanner task performance

Repeated measures analyses were performed to compare in-scanner task performance (correct response and omission rates) between the Panda Games, to identify the easiest and most difficult tasks for patients to perform. Mean performance on the Panda Games task battery was also compared to performance on the VG task, to investigate the effect of developmentally-appropriate task design on performance. Mean performance for the Panda Games was calculated based on the language tasks only (PN, PD and LN) to provide a more valid comparison with VG.

4.4 In-scanner movement

Mean and maximum in-scanner movement measures were obtained from the realignment parameters provided during the realignment process in spm8, and are reported for each task.

4.5 Sensitivity and specificity

The number of active voxels within regions of interest (ROIs) were calculated when u=2.5. At the first level, this threshold is equivalent to a statistical threshold of ~p<0.01 uncorrected. As part of general clinical practice at GOSH, language fMRI results for patients are typically viewed across a range of thresholds from p<0.05 – 0.001 uncorrected. A threshold of ~p<0.01 therefore provides a good estimate for sensitivity to detect activation on an individual basis, at a clinically acceptable and useful level of significance. Using this data I calculated the number and proportion

of participants showing significant activation (> 10 voxels) in each target language region, as one measure of sensitivity. I then calculated the sensitivity and specificity of the Panda Games and VG based on the extent to which target ROIs were activated versus a control ROI. These calculations were based on the formula below:

$$Sensitivity = true\ positives / true\ positives + false\ negatives$$

$$Specificity = true\ negatives / true\ negatives + false\ positives$$

Where:
True positive = significant voxel in the target ROI
False negative = insignificant voxel in the target ROI
True negative = insignificant voxel in the control ROI
False positive = significant voxel in the control ROI

Target ROIs were created using the Automatic Anatomic Labels in the Wake Forest University Pick Atlas using spm8, according to classic language regions; Broca's area (inferior frontal gyrus), Wernicke's area (superior and middle temporal gyri) and Geschwind's territory (angular and supramarginal gyri). I also created target ROIs based on normative data for the Panda Games (Chapter 5) which highlighted the importance of ventral temporal regions in naturalistic use of language, including the temporal pole (middle and superior temporal poles) and a basal temporal region (the inferior temporal gyrus and fusiform gyrus). Intracranial mapping studies have demonstrated the importance of basal temporal regions to naming and comprehension (Davies et al., 1994; Krauss et al., 1996; Luders et al., 1991), which are often targets for epilepsy surgery (Figure 51). Due to the potential for language reorganisation in children with epilepsy (Rosenberger et al., 2009), all target ROIs included both the left and right homologues of each area. All target ROIs are displayed in Chapter 8, Figure 59. The superior parietal lobe, as a region not specified in the functional neuroanatomical model of language (Price, 2012), was used as the control ROI. Two repeated measures analyses were performed to investigate the effect of task on the number of significant voxels, sensitivity and specificity. The first analyses investigated the effect of task within the Panda Games

battery (PN, PD or LN), the second investigated the effect of task when comparing the Panda Games (on average) to VG. The average number of significant voxels, sensitivity and specificity for the Panda Games task battery was based on results from each task when compared to its low-level baseline condition (rest), providing a valid comparison to results from VG (also based on a low-level contrast).

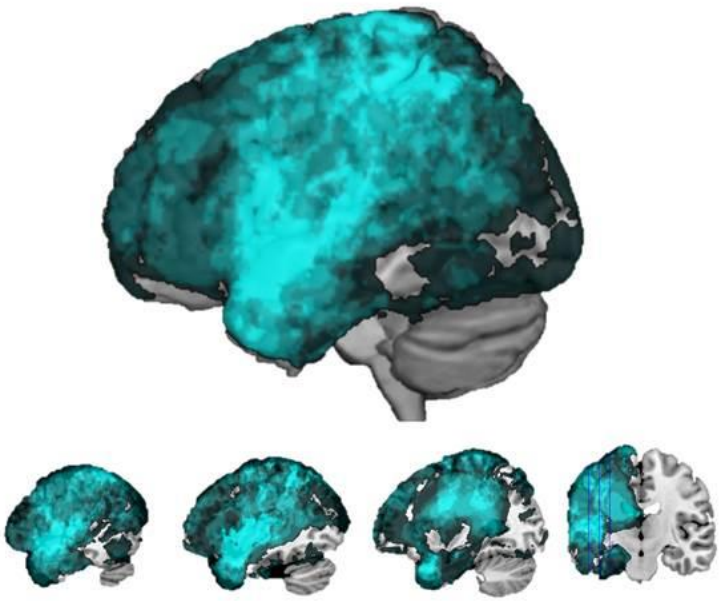


Figure 51: Target regions for neurosurgical resection in children and adolescents with drug resistant epilepsy. Structural lesions are often the target for epilepsy surgery. This lesion overlay map shows the degree of overlap for structural abnormalities in the left hemisphere for a representative sample of surgical candidates. Lighter colours indicate greater overlap between subjects. Lesion data are taken from children and adolescents with drug-resistant epilepsy referred for language fMRI as part of the pre-surgical decision-making process at Great Ormond Street Hospital between 2009 and 2011 (*N*=52, 25 females, mean age = 12.5 +/-2.7 years, mean age of onset of epilepsy = 5.5 +/-3.8). This figure was created using the npm extension for MRIcron (http://www.mccauslandcenter.sc.edu/mricro/mricro/index.html).

5. RESULTS

5.1 In-scanner task performance

Performance data for the group are shown in Table 21, and for individual patients in Table 22. The number of participants performing poorly (<50% correct) for each task was 0 (PN), 1 (PD), 5 (LN) and 2 (VG). An additional three participants achieved only 50% accuracy on LN, suggesting children with epilepsy particularly struggle to perform this task. For those who performed poorly on the LN task, the majority of responses were omissions (rather than errors). This could indicate a slower pace is required to allow patients extra time to process the stimulus and think of the answer. In comparison, common errors on the VG task included repeating the stimulus or generating a noun or adjective, suggesting participants may have misunderstood/forgotten task instructions, or may find generating verbs specifically challenging. For the baseline conditions, three participants performed poorly (<50% correct) on the AG task, suggesting listening tasks in general (under noisy conditions) may be difficult for children with epilepsy. All participants performed CNw and CNs with accuracy >50%.

5.1.1 Correct response rate

Within the Panda Games battery there was a significant effect of task ($F(2,16)=13.47$, $p<0.001$), such that performance was highest for PN, compared to PD ($t(8)=5.53$, $p=0.001$) and LN ($t(12)=4.33$, $p=0.003$). Performance accuracy did not differ between the Panda Games task battery (on average) and VG ($p=0.24$). However, I hypothesised performance differences would be evident when comparing VG to an equivalent (single word generation) task from the Panda Games. For this reason, I performed paired t-tests to compare performance accuracy for VG and PN. Encouragingly, this comparison showed performance accuracy was significantly higher for PN than VG ($t(8)=3.30$, $p=0.01$). Performance did not differ when comparing VG to sentence-level tasks (PD and LN; p values >0.33).

5.1.2 Omission rate

Similar to the correct response rate, there was a significant effect of task on omission rates ($F(2,16)=14.49$, $p<0.001$), such that the omission rates were lower for PN compared to both PD ($t(8)=-2.91$, $p=0.02$) and LN ($t(8)=-4.76$, $p=0.001$). Omission rates were also lower for PD compared to LN ($t(8)=-3.04$, $p=0.02$). There was no significant difference in omission rates between the Panda Games (on average) and VG ($p=0.68$). On an individual-task basis however, there were more omissions during VG than during PN ($t(8)=-4.78$, $p=0.001$), and fewer omissions during VG than LN ($t(8)=-2.89$, $p=0.02$).

5.2 In-scanner movement

In-scanner movement (in millimetres) is shown for each task in Table 21. The size of the maximum movement in any direction during scanning (e.g. head jerks) did not differ between the Panda Games and VG (p values >0.87). Mean movement across the scanning period was also equivalent for the Panda Games and VG (p values >0.52).

	Range	Mean	SD
Practice time (*mins*)			
Panda Games (mean)	0 - 3.92	**0.69**	1.1
Verb generation	0 - 11.37	**1.63**	3
In-scanner task performance			
Correct (%)			
Panda Games (mean)	48 - 92%	**69%**	15%
Picture naming	67 - 100%	**87%**	11%
Picture describing	29 - 92%	**68%**	19%
Listen and name	0 - 92%	**51%**	29%
Verb generation	6 - 97%	**62%**	26%
Omissions (%)			
Panda Games (mean)	3 - 48%	**22%**	14%
Picture naming	0 - 33%	**11%**	12%
Picture describing	4 - 42%	**13%**	13%
Listen and name	4 - 92%	**36%**	26%
Verb generation	0 - 38%	**19%**	12%
Max. movement (*mm*)			
Panda Games (mean)	0.96 - 16.87	**5.88**	5.40
Picture naming	0.57 - 9.30	**2.92**	2.85
Verb generation	0.85 - 6.60	**2.51**	2.20
Mean movement (*mm*)			
Panda Games (mean)	0.09 - 0.61	**0.3**	0.22
Picture naming	0.07 - 0.59	**0.25**	0.19
Verb generation	0.11 - 0.70	**0.29**	0.21

Table 21: Practice time and in-scanner performance data for 13 children with drug-resistant epilepsy. Data for the verb generation task is shown in grey for comparative purposes. Mean practice time (minutes) across all of the Panda Games is shown (this includes practice time on baseline tasks also). Mean performance on the Panda Games represents average performance on the language tasks within the Panda Games task battery only. In-scanner movement is shown in millimetres (mm) for the Panda Games on average (when including baseline tasks) and also for Picture Naming specifically; as the task most comparable to verb generation (i.e. involving single-word generation).

Clinical variables						% correct answers			
Age	**Sex**	**Onset (yrs)**	**Seizure freq.**	**MRI+**	**VIQ**	**PN**	**PD**	**LN**	**VG**
6	F	3	Daily	Yes	93	88%		8%	
6	M	2.5	Weekly	Yes	85	67%			69%
7	F	1	Monthly	Yes	83	88%	63%	46%	78%
8	F	4	Weekly	Yes	79	96%	63%	50%	63%
8	F	4	Daily	Yes	91	96%	67%	29%	6%
9	F	5	Daily	No	79	79%		50%	63%
9	M	2	Weekly	Yes	-	96%	58%	42%	19%
10	M	3	Daily	Yes	60	83%	67%	0%	63%
13	M	2.5	Weekly	Yes	79	100%	92%	83%	72%
13	F	11	Daily	No	-	96%	88%	92%	63%
14	F	1	Daily	No	11	79%		79%	91%
14	M	11	Weekly	Yes	108	71%	29%	50%	59%
16	F	5	Monthly	No	57	96%	83%	83%	97%

Table 22: Task performance case summaries for children with epilepsy. Task performance data are presented alongside clinical variables, including: age, sex, age of seizure onset (years), seizure frequency, MRI findings and Verbal IQ (VIQ). Tasks which were not performed due to scanning are in black. Tasks where performance accuracy =< 50% are highlighted in grey. Picture Naming (PN), Picture Describing (PD), Listen and Name (LN) and verb generation (VG). Two patients did not have neuropsychological data (-).

5.3 Sensitivity and specificity

The number of participants showing significant activation in language regions (when $p<0.01$ and $k>01$) did not differ between tasks (Table 23). Neither sensitivity nor specificity differed between the Panda Games (PN, PD and LN). Further, there was no significant difference in sensitivity or specificity between the Panda Games (on average) and VG (Table 23). There was no significant interaction of task and region.

	Panda Games																							
	PN			PN>CNw			PD			PD>CNs			LN			LN>AG			MEAN			VG		
Activation detected	*N*	%		*N*	%		*N*	%		*N*	%		*N*	%		*N*	%		*N*	%		*N*	%	
Broca	7	54%	-	6	46%	-	6	67%	-	7	78%	-	5	42%	-	6	50%	-	11	85%	-	8	67%	-
Wernicke	9	69%	-	11	85%	-	6	67%	-	6	67%	-	11	92%	-	7	58%	-	13	100%	-	11	92%	-
Geschwind	3	23%	-	2	15%	-	3	33%	-	2	22%	-	0	0%	-	2	17%	-	6	46%	-	4	33%	-
Temporal pole	4	31%	-	4	31%	-	4	44%	-	3	33%	-	6	50%	-	4	33%	-	8	62%	-	7	58%	-
Basal temporal	11	85%	-	13	100%	-	8	89%	-	8	89%	-	5	42%	-	7	58%	-	13	100%	-	9	75%	-
Sensitivity	**Max.**	**Mean**	**Std**	**Max.**	**Mean**	**Std**	**Max.**	**Mean**	**Std**	**Max.**	**Mean**	**Std**	**Max.**	**Mean**	**Std**	**Max.**	**Mean**	**Std**	**Max.**	**Mean**	**Std**	**Max.**	**Mean**	**Std**
Broca	0.11	0.03	0.04	0.06	0.01	0.02	0.18	0.04	0.06	0.12	0.04	0.05	0.23	0.04	0.08	0.23	0.04	0.07	0.12	**0.04**	0.04	0.06	**0.03**	0.03
Wernicke	0.29	0.04	0.08	0.23	0.03	0.06	0.18	0.04	0.06	0.14	0.03	0.05	0.23	0.08	0.07	0.09	0.02	0.03	0.12	**0.05**	0.04	0.15	**0.05**	0.05
Geschwind	0.59	0.05	0.16	0.52	0.04	0.14	0.56	0.10	0.20	0.46	0.08	0.16	0.01	0.00	0.00	0.07	0.01	0.02	0.20	**0.04**	0.07	0.12	**0.02**	0.03
Temporal pole	0.06	0.01	0.02	0.07	0.01	0.02	0.14	0.04	0.05	0.07	0.02	0.03	0.10	0.02	0.03	0.10	0.03	0.04	0.06	**0.02**	0.02	0.14	**0.03**	0.04
Basal temporal	0.33	0.10	0.10	0.24	0.05	0.07	0.21	0.08	0.08	0.15	0.04	0.05	0.22	0.03	0.07	0.22	0.03	0.06	0.11	**0.07**	0.04	0.20	**0.03**	0.05
Specificity																								
Broca	1.00	0.92	0.25	1.00	0.93	0.24	1.00	0.92	0.13	1.00	0.95	0.10	1.00	0.98	0.06	1.00	0.98	0.06	1.00	**0.94**	0.09	1.00	**0.93**	0.14
Wernicke	1.00	0.92	0.25	1.00	0.93	0.24	1.00	0.92	0.13	1.00	0.95	0.10	1.00	0.98	0.06	1.00	0.98	0.06	1.00	**0.94**	0.09	1.00	**0.93**	0.14
Geschwind	1.00	0.92	0.25	1.00	0.93	0.24	1.00	0.92	0.13	1.00	0.95	0.10	1.00	0.98	0.06	1.00	0.98	0.06	1.00	**0.94**	0.09	1.00	**0.93**	0.14
Temporal pole	1.00	0.92	0.25	1.00	0.93	0.24	1.00	0.92	0.13	1.00	0.95	0.10	1.00	0.98	0.06	1.00	0.98	0.06	1.00	**0.94**	0.09	1.00	**0.93**	0.14
Basal temporal	1.00	0.92	0.25	1.00	0.93	0.24	1.00	0.92	0.13	1.00	0.95	0.10	1.00	0.98	0.06	1.00	0.98	0.06	1.00	**0.94**	0.09	1.00	**0.93**	0.14

Table 23: Sensitivity and specificity of the Panda Games task battery compared to a task currently used in the pre-surgical setting (verb generation; VG). The number and proportion of participants in which activation was detected within each language region ($p<0.01$ and $k>10$) are shown for each task. Sensitivity and specificity were calculated as the extent of activation in language regions versus a non-language control region (the superior parietal lobe). The maximum (max.) observed in the sample and mean values are shown for each task in the Panda Games battery, as well as the mean for the entire Panda Games battery (when tasks are compared to a rest baseline). Data are also shown for verb generation (VG), which is currently used for assessing language in the pre-surgical setting. Sensitivity and specificity were calculated when $u=2.5$ (which corresponds to a statistical threshold of approximately $p<0.01$ uncorrected at the first-level).

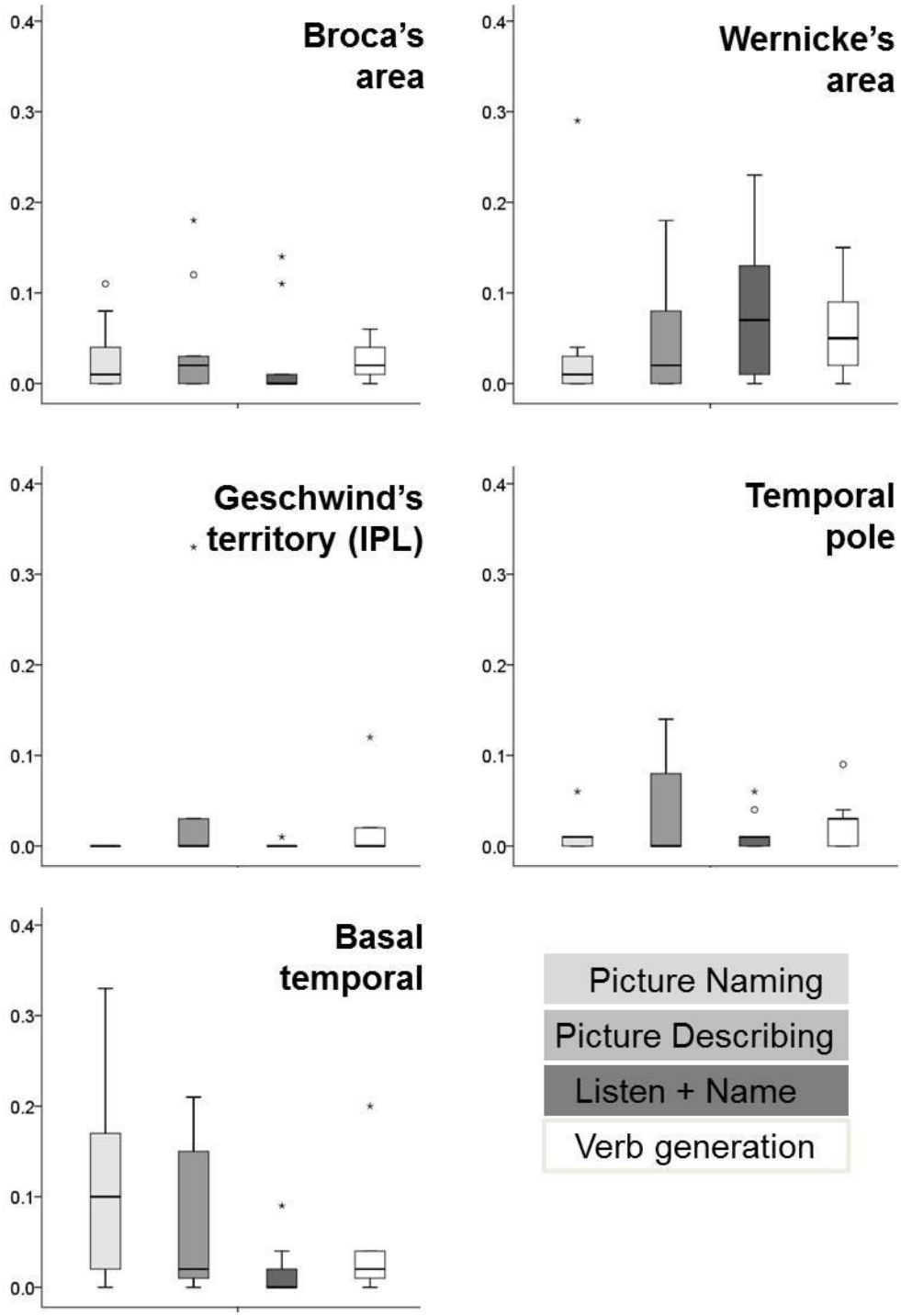


Figure 52: Between-task differences in the sensitivity to detect activation in language regions. Here sensitivity is calculated based on the extent of activation in language regions versus a non-linguistic control region. Outliers are indicated.

5.4 Patient Case studies

Here I report two case studies, to provide a more in depth investigation into the contribution of the new Panda Games task battery to the pre-surgical decision-making process. These case studies demonstrate the challenges of performing f MRI in children with cognitive impairment (Patient F) and behavioural co-morbidities (Patient A), as well as highlighting important issues regarding the interpretation and validation of fMRI results. Finally, findings from these studies raise important issues which I discuss further in terms of study limitations and directions for future work at the end of this thesis.

5.5 Case 1: Patient A

Patient A is a 6 year old female with drug resistant symptomatic focal epilepsy, who experienced her first seizure at the age of three years. Typically patient A experiences two clusters of asymmetric spasms (with some atonic posturing) per day, lasting up to ~3 minutes. She was taking one antiepileptic medication at the time of testing (levetiracetam), and was trialling a homeopathic medicine. Video telemetry provided strong support for a left sided origin for interictal and ictal events, and structural MRI showed a left temporal porencephalic cyst (thought to be the result of perinatal intraparenchymal haemorrhage) (Figure 53A). Patient A performed within the normal range on standardised neuropsychological assessments of verbal and nonverbal IQ (WPPSI 3rd edition) and academic attainment (WIAT-II), but concerns regarding hyperactivity had been highlighted by parents and teachers, although a formal diagnosis of ADHD had not been made at the time she was seen at the clinic. Patient A underwent language fMRI to identify the language dominant hemisphere, and intracranial mapping (ICM) to localise language regions surrounding the site of the lesion. She also applied to take part in the Panda Games study. Because stimuli from the Panda Games battery were also used during ICM, this patient case study provides an example of fMRI validation with invasive measures.

5.5.1 Language fMRI

Patient A practised VG and the Panda Games at home prior to her appointment, and was looking forward to playing the Panda Games. She showed good understanding of task instructions prior to scanning.

5.5.1.1 Standard clinical protocol

During scanning, verb generation had to be aborted; she found this task challenging inside the scanner and was very distressed by the white noise stimuli used in the baseline condition.

5.5.1.2 Panda Games

Patient A played the Panda Games after the standard clinical protocol had been completed. She performed PN (88% correct) and CNw (71%) well, and found the picture stimuli very engaging, but fell asleep during PD. During LN Patient A only responded on 8% of trials, and did not respond on any trials for AG. She may have been attending to stimuli however (passively). When asked why she didn't respond, Patient A said she didn't want to; at this point scanning was aborted. Functional MRI results are shown on a rendered image of the brain in Figure 53B for PN (compared to CNw; to localise core linguistic processing). Results for LN were of an unusable quality, due to non-compliance as reported above, and are not shown.

5.5.2 Intracranial mapping methods

Two strips and three depth electrodes were implanted (Figure 53A and Table 24). Electrodes were stimulated using the Penfield protocol (Penfield & Boldrey, 1937), with a pulse duration of 500 for subdural electrodes and 1000 for depth contacts. Stimulus intensities ranged from 3.0-15.0 mA for subdural and 1.0-3.0 for depth contacts. During stimulation, Patient A played the Listen and Name game (naming from auditory definitions). The Listen and Name game was used to distinguish between articulatory/production errors (delayed naming) and comprehension errors (failed naming).

5.5.2.1 Limitations to this validation procedure

There are several limitations to the validation of fMRI with ICM, which I outline here before presenting results from Patient A. First, ICM itself provides only approximate results; stimulation can have local and distal effects beyond the electrode under investigation. Second, like fMRI, ICM is a threshold-dependent technique, with the strength of stimulation effecting the severity and extent of disruption to cortical function. Third, identification of language sites with ICM is limited by the placement of the electrode grid. By identifying only a selection of language sites, ICM may overestimate the false positive rate of fMRI (which identifies distributed language regions throughout the whole brain). Finally, ICM is very challenging in children, and especially those with cognitive impairment or behavioural co-morbidities. For both ICM and fMRI, there are time limitations on the amount of mapping that can be performed, as well as co-operation issues. Consequently, obtaining good quality data in both modalities is not always possible. For example, Patient A was engaged by PN but not LN during fMRI scanning, but was engaged by LN and not PN during ICM. As a result, it was not possible to validate fMRI results with ICM using the same task (LN data from fMRI was unusable, and PN was aborted during ICM). This is the main limitation of most fMRI validation studies reviewed in the introduction to this thesis (section **Error! Reference source not found.**). Previous findings, and my findings reported here, may therefore underestimate the sensitivity and specificity of fMRI for identifying critical language regions.

5.5.3 Positive language sites on ICM

Positive language sites were identified on the left anterior strip (electrodes 5 and 6; highlighted in red on Figure 53A and C). Stimulation here caused articulatory errors during naming (delayed naming and lip movements), consistent with their location in proximity to motor cortex. Positive language sites were also identified on the left posterior superior temporal strip (electrodes 2-5; highlighted in red on Figure 53A and D). Stimulation of electrodes 2 and 5 produced delayed naming. Once the stimulation ended at

electrodes 2 and 5 Patient A was able to produce the correct name, suggesting a production error induced by activation of these sites. Stimulation of electrodes 3 and 4 produced comprehension errors; Patient A was unable to name the item even once stimulation ended, at which point she responded "I can't hear you", suggesting interrupted speech comprehension during stimulation. Electrodes 3 and 4 were located medially within the superior temporal plane, and into the white matter (Figure 53D; electrode 3 is located to the right of the cross-hair). In total there were six positive sites found during ICM; four associated with production/articulatory errors, two associated with semantic processing (Table 24).

5.5.4 Sensitivity and specificity of fMRI

Functional MRI identified widespread activation in the left temporal pole, posterior inferior temporal gyrus and posterior middle temporal gyrus during picture naming (Figure 53B). This highlighted active tissue in proximity to the anterior and posterior edges of the lesion (not shown). The anterior cluster included the left parahippocampal gyrus and hippocampus. The posterior cluster was located in the fusiform gyrus/lingual gyrus. These regions were not tested with ICM, so it is not possible to verify whether these reflect true or false positives. However, they emphasise the potential for fMRI to identify *networks* supporting cognition, whereas ICM can only identify localised activity in selected regions (limited by the size of the strip, grid or depth electrode). The true/false positive rates and true/false negative rates for language fMRI in Patient A are displayed in Table 24. In total fMRI showed a sensitivity of 0.50, and specificity 0.70. However, two positive sites from ICM were located in the white matter, which is not assessed with fMRI. When the true positive rate was limited to grey matter sites, the sensitivity of fMRI increased to 0.75. True positives and false positives identified on fMRI are described below, and examples are shown in Figure 53C-E.

5.5.4.1 True positives

Using the contrast PN>rest (which identifies the whole picture naming network, including articulation), fMRI identified true positive activation

located around electrode 5 on the left anterior strip (Figure 53C). Activation in this region was not apparent when comparing PN to CNw (which controls for articulatory processing). Functional MRI findings were therefore consistent with findings from ICM; stimulation of this region induced articulatory errors. Using the contrast PN>CNw (which aims to localise linguistic processing during picture naming), fMRI identified true positive activation located between electrodes 2-3 in the left posterior superior temporal gyrus depth (Figure 53D). Stimulation of electrode 2 during ICM was associated with production errors (delayed naming), whereas stimulation of electrode 3 was associated with comprehension errors, supporting the involvement of these regions in linguistic processing. Crucially, activation in these sites was only apparent for fMRI at a low statistical threshold (p<0.05 uncorrected).

5.5.4.2 False positives

Functional MRI identified false positives on the inferior posterior temporal depth and hippocampal depth (the inferior temporal depth is shown as an example in Figure 53E). Ictal and interictal activity were recorded across all sites on both these depth electrodes during ICM. Whether or not epileptiform activity in these regions induced activation of equivalent sites on fMRI is not known. However, combined EEG-fMRI holds the potential to improve the sensitivity of pre-surgical language fMRI by controlling for non-cognitive (epileptiform) activity during fMRI (Chaudhary et al., 2009; Thornton et al., 2009; Vulliemoz et al., 2009), and potentially reducing the false positive rate of MRI. Functional MRI also showed one false positive activation locus on the posterior superior temporal depth, at electrode 1 (toward the lateral surface). Electrodes 2-5 were implicated in language processing on ICM, however electrode 1 did not cause production or comprehension errors on stimulation. Functional MRI activation at the site of electrode 1 could be a result of statistical smoothing during data pre-processing, which aims to improve the signal-to-noise ratio, but also reduces spatial precision. Functional MRI activation at the location of electrode 2 (a true positive site) may therefore have spread to encompass electrode 1 during the smoothing process. Whether or not

smoothing is advantageous for pre-surgical fMRI has not been formally tested, however it is likely to improve sensitivity but reduce specificity (as can be seen here). No false positives were identified on the left anterior strip or the left anterior superior temporal depth.

5.5.5 Summary: Patient A

For Patient A, fMRI performed using the Panda Games achieved a sensitivity of 0.50, suggesting accurate identification of 50% of critical language sites. This is relatively low compared to previous studies, and could result in surgical removal of eloquent tissue. However, two of the true positive sites identified with ICM were located within (or on the border of) white matter, which is not identified with fMRI. When restricting the calculation of sensitivity to only grey matter sites, fMRI achieved a sensitivity of 0.75, which is in line with previous studies (sensitivity of ~0.59-1.00; reviewed in section 1.2 of the introduction to this thesis). This finding highlights a limitation of fMRI for identifying critical regions (some critical regions lie in the white matter!) and emphasises the need for supplementary information from diffusion tensor imaging (DTI) in pre-surgical language mapping. Another important issue raised by Patient A, is the effect of statistical thresholding in the pre-surgical setting. The 'critical' activation identified in the left superior posterior depth (Figure 53D) was seen on fMRI at a low statistical threshold (p<0.05 uncorrected), but not at a higher threshold (p<0.01 uncorrected); even though this statistical threshold is still considered very liberal in the neuroimaging field. False negatives could have fatal consequences in the pre-surgical setting, potentially resulting in removal of eloquent cortex. This emphasises the need for extensive validation of fMRI with invasive recordings, to identify methods which can determine the appropriate statistical threshold to reduce false negatives and false positives, on an individual basis.

My results for Patient A showed fMRI had a greater specificity to detect critical language sites (0.70) relative to these previous studies (~0.54) (reviewed in section 1.2 of the introduction to this thesis). My finding of a relatively high

specificity for fMRI may have been due to use of a high-level baseline task (which controls for sensory-motor aspects of task performance), developmental-matching of stimuli (which may reduce identification of non-critical regions associated with sustaining effortful task performance), or event-related analysis (which may model the BOLD response more accurately). The effects of each of these variables on the specificity of fMRI requires formal testing in relation to ICM.

Figure 53: (Figure overleaf) Validation of functional MRI with intracranial mapping in a 6 year old girl (Patient A).

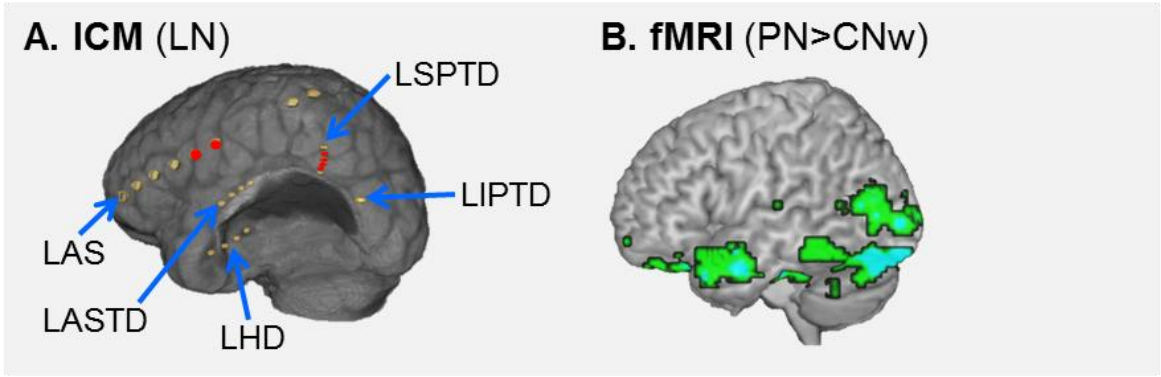
A) Placement of the left anterior strip (LAS), left anterior superior temporal depth (LASTD), left hippocampal depth (LHD), left inferior posterior temporal depth (LIPTD), and a left superior posterior temporal depth (LSPTD) are shown in yellow. Electrodes associated with language disruption during performance of the Listen and Name game (LN) during intracranial mapping (ICM) are highlight in red.

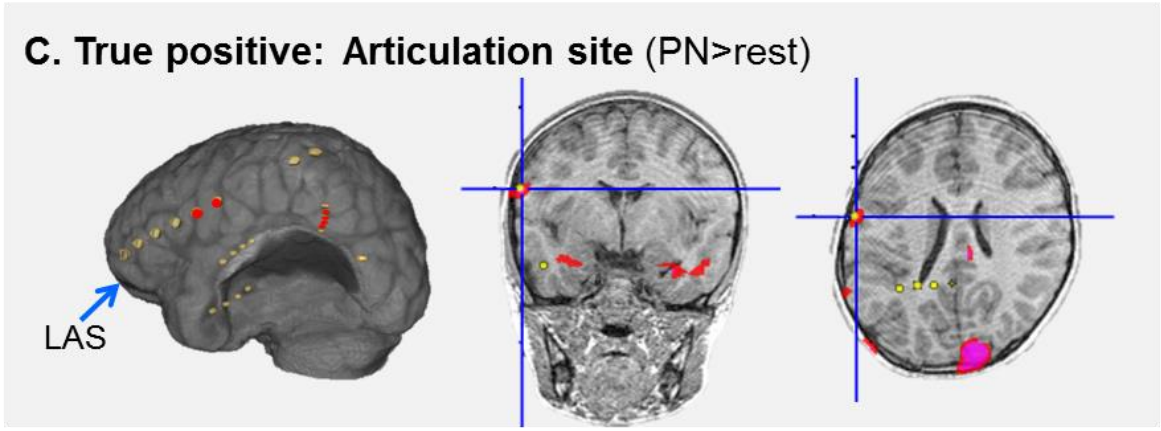
B) Activation on fMRI during performance of the Picture Naming game from the Panda Games task battery (when compared to its sensory-motor baseline task; Colour Naming with words). All fMRI activations for this contrast are shown in cyan ($p<0.01$ uncorrected) and green ($p<0.05$ uncorrected).

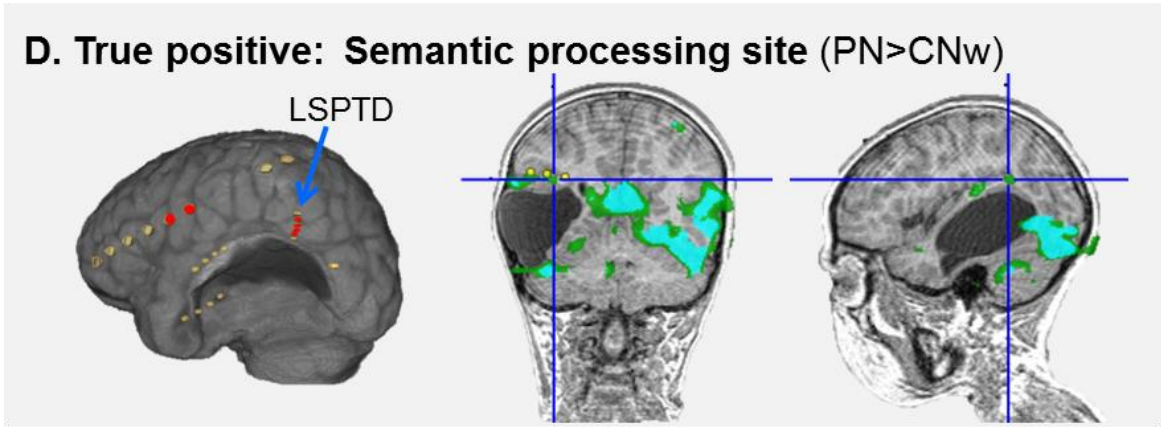
C) True positive fMRI activation associated with articulation on both fMRI and ICM at electrode 5 on the left anterior strip (red=$p<0.01$, pink=$p<0.05$).

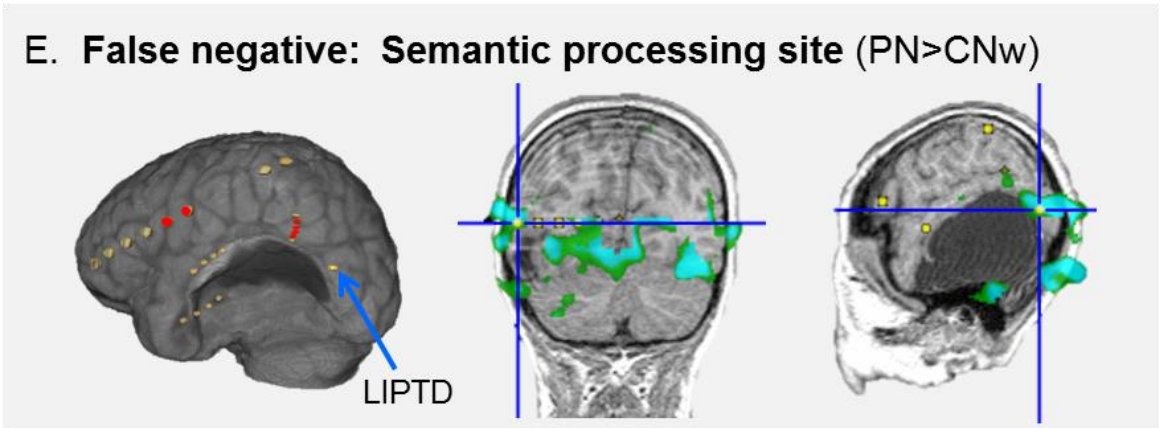
D) True positive fMRI activation associated with core linguistic processing on both fMRI and ICM, located between electrodes 2 and 3 on the superior posterior temporal depth (shown by the cross-hairs). Note this 'critical' activation is only apparent when $p<0.05$ uncorrected.

E) False negative fMRI activation located at electrode 1 of the inferior posterior temporal depth; a region implicated in the ictal onset of seizures which may or may not have been active during fMRI. This activation could reflect core linguistic processing or epileptiform activity during fMRI.


A. ICM (LN)
LSPTD
LIPTD
LAS
LASTD
LHD
B. fMRI (PN>CNw)


C. True positive: Articulation site (PN>rest)
LAS


D. True positive: Semantic processing site (PN>CNw)
LSPTD


E. False negative: Semantic processing site (PN>CNw)
LIPTD

Intracranial mapping results			Functional MRI results			
Electrode	**Language disruption**	**No. contacts**	**TP**	**FP**	**TN**	**FN**
Left anterior strip (LAS)	Articulatory errors (5+6)	6	1	0	4	1
Left anterior superior temporal depth (LASTD)	No disruption	4	0	0	4	0
Left hippocampal depth (LHD)	No disruption	4	0	2	2	0
Left superior posterior temporal depth (LSPTD)	Articulatory errors (2+5) Comprehension errors (3+4)	6	2	1	1	2
Left inferior posterior temporal depth (LIPTD)	No disruption	6	0	3	3	0
	TOTALS:	**26**	**3**	**6**	**14**	**3**

Sensitivity:	**0.50**
Specificity:	**0.70**

Table 24: Sensitivity and Specificity of fMRI validated with intracranial mapping (ICM) in a 6 year old girl (Patient A). The type of language disruption observed for each electrode during ICM is described. The number of true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN) found with fMRI are listed for each electrode. The overall sensitivity and specificity of fMRI is shown. Note; when two white matter true positive sites are discounted, the sensitivity of fMRI for detecting language sites in the grey matter is 0.75.

5.6 Case 2: Patient F

Patient F is a 10 year old male with drug resistant focal epilepsy, who experienced his first seizure at three years of age. Typically Patient F experiences 2-3 focal seizures per day, which include visual aura, oral and hand automatisms, and speech arrest. Video telemetry provided evidence for ictal and interictal EEG activity in the left posterior quadrant, consistent with findings of a dysembryoplastic neuroepithelial tumour (DNET) in the left posterior occipito-temporal cortex (Figure 54). This lesion was located in a region of the left occipito-temporal cortex which included the visual word form area and slightly more anterior fusiform regions associated with amodal semantic processing, which are discussed at length in Chapters 5 and 8. According to standardised neuropsychological assessment using the WISC-IV, Patient F had significant cognitive impairment (General Ability Index=63), including poor verbal comprehension (standard score=60) and working memory (standard score =55). His nonverbal abilities were significantly stronger than his verbal abilities (perceptual reasoning index=72). Patient F underwent language fMRI to identify the language dominant hemisphere, and also applied to

take part in the Panda Games study. The surgeon specifically required information about lateralisation of temporal activation, to indicate the risk of temporal lesionectomy to language function.

5.6.1 Language fMRI

Patient F practiced the Panda Games and VG at home prior to his appointment, and also underwent mock scanning. He showed good understanding of task instructions prior to scanning, but was extremely shy and was distressed when he did not know the correct answer.

5.6.1.1 Standard clinical protocol

Patient F performed well on the VG task (63% correct) but was only able to answer 2/8 comprehension questions correctly following the passive story comprehension task, in line with his low verbal comprehension and working memory scores on neuropsychological assessment.

5.6.1.2 Panda Games

Based on his neuropsychological profile, Patient F performed level 1 of the Panda Games (5-7 year old level). Patient F performed well on PN (83% correct) and PD (67%), despite cognitive impairment. However, for LN he had an error rate of 29%, and failed to produce any response for the remaining 71% of trials. The error responses suggest Patient F forgot the task instructions during scanning (in line with his very poor verbal working memory), and instead attempted to generate verbs similar to the auditory task he had performed earlier. For example, when he heard the 'fridge' stimulus ("A cupboard that keeps food cold") he responded "to eat", and when he heard the 'toilet' stimulus ("Something in the bathroom that flushes") he responded "flush the water". At this point – and likely due to distress caused by poor performance – Patient F asked to finish scanning.

5.6.2 Summary: Patient F

Strong activation was seen in the unaffected right occipito-temporal cortex during both PN and PD, but not VG (Figure 54B). Picture describing produced strong activation along basal regions of the right temporal lobe, extending to the anterior

temporal pole and the inferior frontal gyrus (BA 45/47). Laterality indices (Table 25) in Wernicke's area (superior and middle temporal gyri) and the basal temporal region (inferior temporal and fusiform gyri) were strongly right lateralised for PN and PD, but were strongly left lateralised for VG (which showed little activation in these regions). Activation in Broca's area (inferior frontal gyrus) was bilateral (PN) or strongly right lateralised (PD) for the Panda Games, but was left lateralised for VG (again, which showed little activation in these regions). These findings could provide evidence for a dissociable effect of input modality (pictorial versus auditory) on lateralisation of activation in the language network. However, the quality of the VG scan was poor compared to the typical activation pattern seen in patients. As such, it is not possible to distinguish between the effects of data quality and input modality on these findings. They do however, emphasise the importance of performing multiple runs of language fMRI to guarantee at least some good quality data is obtained, and to allow validation of findings across modalities, tasks and across repeated runs.

	Laterality Indices		
Region	**PN**	**PD**	**VG**
Broca	-0.15	-0.76	0.37
Wernicke	-0.67	-0.66	0.76
Basal temporal	-0.62	-0.38	0.61

Table 25: Laterality indices in Broca's area, Wernicke's area, and the basal temporal language area across three language tasks in a 10 year old boy (Patient F). Laterality indices are shown for Picture Naming (PN), Picture Describing (PD) and Verb Generation (VG), when each task is compared to its sensory-motor baseline task. Laterality indices are shown for Broca's area (inferior frontal gyrus), Wernicke's area (superior and middle temporal gyri) and a basal temporal region being considered for surgical resection in this patient (inferior temporal gyrus and fusiform gyrus).

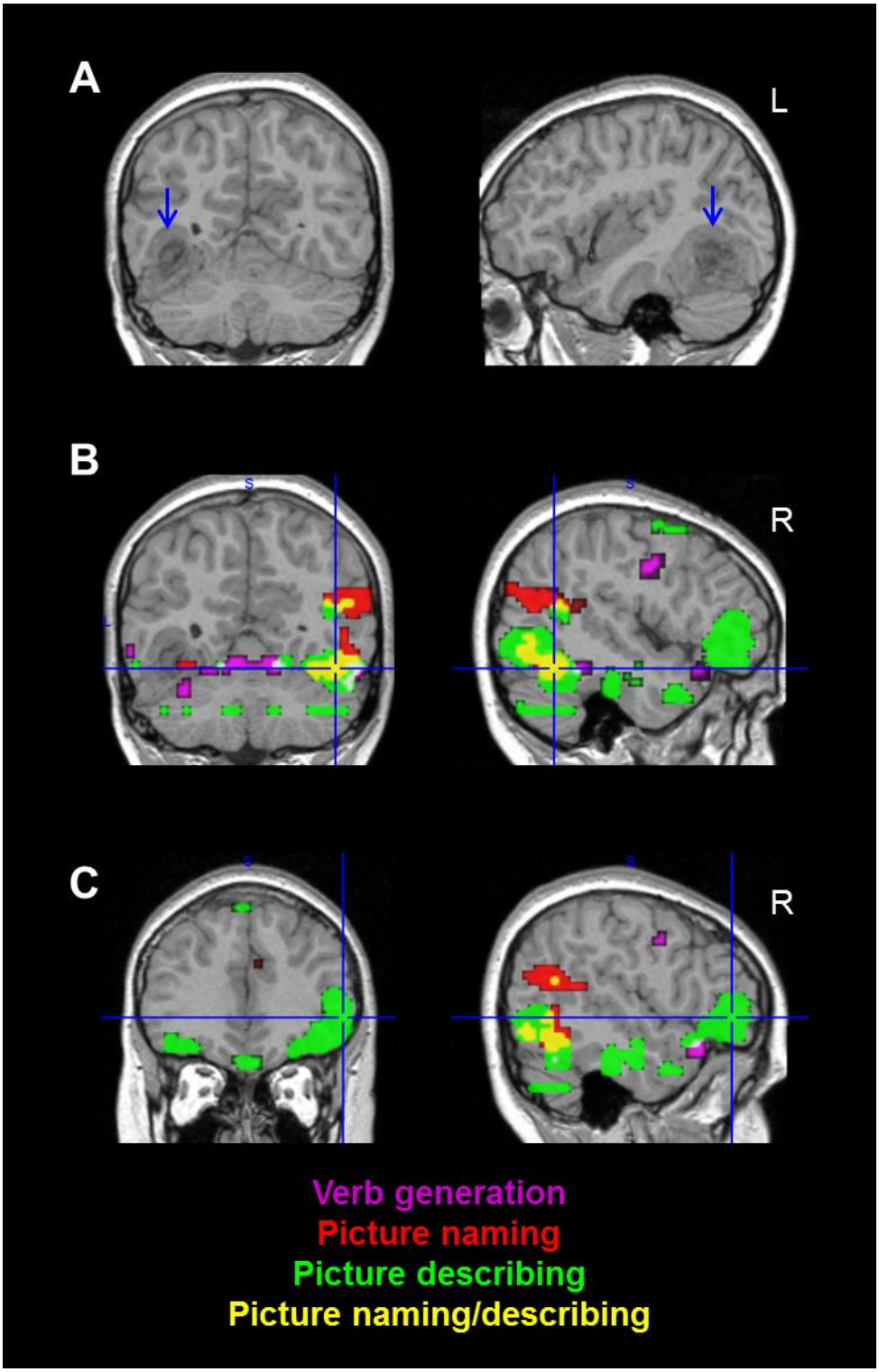


Figure 54: Language mapping with fMRI in a 10 year old boy (Patient F) with a left posterior occipito-temporal dysembryoplastic neuroepithelial tumour (A). Functional MRI results for Verb Generation (pink), Picture Naming (red), Picture Describing (green) and associated with both Picture Naming and Picture Describing (yellow) are

shown at a threshold of $p<0.01$ uncorrected, in the right occipito-temporal region (B) and right inferior frontal gyrus (C).

6. DISCUSSION

Pre-surgical candidates at high risk for post-operative language problems are identified according to several criteria in the pre-surgical setting; disproportionate impairment of verbal abilities (relative to non-verbal abilities), speech disruption during seizures, or structural abnormality in proximity to perisylvian cortex. These patients form the majority of referrals for language fMRI on an epilepsy surgery programme. Pre-surgical language fMRI is therefore most needed for exactly those patients who will find language fMRI most challenging. This situation is amplified in the paediatric population, whose poor verbal abilities may be compounded by immaturity, anxiety or difficult behaviour, and can make language fMRI extremely challenging. The Panda Games have been designed to be child-friendly, with four difficulty levels, such that patients with verbal abilities equivalent to a five year old should be able to undergo language fMRI. Results from healthy children aged 5-16 years suggest the Panda Games do provide successful mapping of receptive and expressive language networks in children of the same age as those referred for language fMRI as part of a paediatric epilepsy surgery programme in the UK (5-16 years). However, due to the cognitive and behavioural profile of children with drug-resistant epilepsy (Bailet & Turk, 2000; Davies, Heyman, & Goodman, 2003; Rutter, Graham, & Yule, 1970), it is important to pilot the new fMRI tasks in this population specifically. It is also important to answer several questions regarding the reliability, sensitivity and usefulness of information obtained from the new Panda Games task battery for pre-surgical planning. Results from this pilot study are discussed below to answer such questions, and are discussed in relation to results from a verb generation paradigm, which is currently (and commonly) used to assess language lateralisation in the pre-surgical setting, and provides a valuable point of comparison.

6.1 Is it possible to obtain usable (good quality) scans with the Panda Games?

6.1.1 In-scanner task performance

The age-appropriate task design in the Panda Games task battery aimed to improve task performance accuracy. However, contrary to hypothesis task performance was equivalent for the Panda Games and for VG (which provides a one-size-fits-all approach to task design). Comparing average performance on the Panda Games and VG may not be a valid comparison however, as tasks in the Panda Games battery require sentence-level language processing while VG is a word generation task. Indeed when word-level tasks were compared, performance accuracy was significantly higher for the Panda Games task (PN) relative to VG. Of the Panda Games which were performed, patients found LN the most difficult to perform, with the majority performing ~50% correct. The high omission rate for this task (36%) – and relatively low error rate (13%) – could suggest this task was too fast-paced for some patients, who may have struggled to find and produce an answer within four seconds. In contrast, although performance accuracy was equivalent for LN and VG, VG showed equal amounts of errors (19%) and omissions (19%). Errors typically included repetition of the stimulus or generation of a noun, suggesting generating verbs is particularly challenging for some patients. With a longer inter-stimulus interval, patients may have achieved significantly higher scores for LN, possibly with performance accuracy exceeding that of VG; although this requires empirical support.

6.1.2 Performance of young patients (aged 5-7 years)

Two participants in this study were aged 6 years, which is below the typical minimum age for pre-surgical language fMRI at GOSH (7 years). Neither patient was able to perform VG; one patient (Patient A) was unable to generate verbs and was distressed by the white noise stimuli, while the second patient could not generate verbs and instead repeated the stimulus for the majority of trials. In contrast, both patients were able to perform PN; achieving 88% and 67% accuracy respectively. These patients either fell asleep or did not comply with further scans however (both had behavioural difficulties). For these patients, language fMRI failed based on VG alone. The Panda Games battery however provided useable

information regarding localisation of the word generation network, which was particularly useful in the case of Patient A. Based on these patients alone it is difficult to determine whether the Panda Games can reliably extend the use of pre-surgical language fMRI to children < 7 years of age. Further investigations are required to determine whether young children with drug-resistant epilepsy who do not have comorbid behavioural difficulties are able to comply with the scanning procedure and perform the Panda Games. For children with challenging behaviour further work is needed to develop a pre-scan preparation protocol which can enhance compliance throughout scanning. In-scanner protocols which aim to sustain engagement and compliance through the use of small immediate rewards may be particularly helpful for these children (Schlund et al., 2011). Anecdotally, patients were more engaged by tasks with picture stimuli (PN and PN) than auditory tasks (LN and VG), suggesting the use of pictures in auditory language tasks may be helpful in enhancing task compliance and performance. The child-friendly design of all task materials was also valuable; many patients came to the appointment prepared and excited to play the Panda Games, and there were no problems with any patients getting into the bore of the magnet. One child aged 5 years (who was excluded from this pilot) had not been shown the Panda Games materials by her parents before her appointment, and did not want to go in the scanner. For this patient more extensive preparation work and introduction to the Panda Games prior to scanning may have been helpful. For patients with cognitive impairment (e.g. Patient F) obtaining information regarding their neuropsychological profile is extremely valuable (although this was not always possible for the patients in this pilot study). For example, specific strategies could be developed to help children with very poor working memory.

6.1.3 In-scanner movement

In-scanner movement did not differ between the Panda Games task battery and VG, suggesting increased articulatory demands in the Panda Games are not associated with poorer data quality. On average, movement was less than 1mm in patients. Instead, large jerks appear more of a problem for this population. My results suggest jerks could be as large as 17mm over the course of a long scanning session. However, post-processing methods can be used to remove volumes

obtained during these brief and infrequent large movements (e.g. scan nulling discussed in Chapter 4).

6.2 Do the Panda Games produce reliable activation in language regions on an individual basis?

The Panda Games were as sensitive as VG in detecting activation in language regions on an individual basis. On average, sensitivity for the Panda Games was highest for Wernicke's area (0.05) and the basal temporal region (0.07). Although not statistically significant, sensitivity to detect activation in basal temporal regions was lower on average for VG (0.03). This measure of sensitivity reflects the proportion statistically significant voxels at a given threshold (here $p<0.01$ uncorrected) in the target language region (true positives), relative to a control region (false positives). The target regions used in this study are large and over-inclusive; to capture individual variability in localisation of language in children with epilepsy (Liegeois et al., 2004). Thus, my results may underestimate the sensitivity of both the Panda Games and VG to detect language activation. For the purposes of between-task comparisons however this remains a valid measure. A more realistic estimate of each task's sensitivity to detect activation can be obtained by looking at the proportion of participants showing any activation in target regions, regardless of extent (except for specifying a minimum cluster size of 10, to control for random noise). My results from this analysis show no significant difference in the proportion of participants with activation in language regions for the Panda Games versus VG. For the Panda Games 100% of participants showed activation in Wernicke's area and the basal temporal region during at least one of the tasks. These results demonstrate the value of using a battery of tasks, as opposed to a single task, for pre-surgical language mapping (Gaillard et al., 2004; Wilke et al., 2010). Although I hypothesised activation in Broca's area for all tasks (Chapter 6), activation was seen in 42-62% of participants across the three Panda Games, suggesting activation in this region is highly variable. Fifteen Percent of participants showed no activation in Broca's area for any of the Panda Games, and 33% showed no activation during VG; which was specifically designed to activate this region. For clinical purposes, a longer run of VG (eight blocks) is typically performed during fMRI scanning. It is

possible that variable activation of Broca's area during VG in this study reflects reduced power to detect activation when a shorter run (four blocks) is used.

6.3 Do the Panda Games produce localised activation in target regions?

All tasks within the battery showed high specificity (mean range=0.92-0.98). Verb generation was also highly specific (mean=0.93), and there were no significant differences in the specificity of the Panda Games versus VG, or between the Panda Games themselves. These findings suggest activation induced during performance of the Panda Games is localised to regions implicated in language, with a low false positive rate. In the pre-surgical setting false positive findings may implicate non-critical tissue in language function, which may affect the prognosis for post-surgical outcome and the decision for surgery. It is therefore encouraging that this new battery of tasks has the same specificity as VG, which has been validated using invasive measures (Arora et al., 2009; Liegeois et al., 2002).

6.4 Do the Panda Games provide new information useful to pre-surgical planning?

Both of the patient case studies presented in this chapter show scenarios in which the Panda Games provide information which is not available when VG is used alone. Most notably, the Picture Naming and Picture Describing tasks provided mapping of posterior occipito-temporal regions which was valuable to both Patients A and F. For Patient F these tasks were particularly useful in showing the clear right lateralisation of function in these regions, away from the affected hemisphere. For Patient A, Picture Naming during fMRI identified active tissue surrounding the lesion, not apparent with VG. Further, when tested with ICM, these regions proved critical to language function. My data suggest temporal regions are the target for surgery in ~40% of cases, similar to a previous multicentre study of surgery for paediatric epilepsy (Harvey et al., 2008). The proportion of patients with more posterior temporal lesions or ictal onset are not yet known, however these initial findings suggest the Panda Games may provide useful information regarding localisation and lateralisation of language function for at least up to 40% of patients.

6.5 Statistical thresholding of fMRI results in the pre-surgical setting

The results obtained from fMRI differ according to which statistical threshold is used. This is discussed at length in Chapter 3 (section 7.1). For experimental investigations of language using group maps, a threshold of $p<0.05$ corrected for family wise error is generally accepted, with some authors also using the more liberal threshold $p<0.001$ and $k>10$ (uncorrected). Clinically however, individual patients' fMRI results are typically viewed across a range of statistical thresholds, to get an impression of which activations are more reliable and which are less robust (usually performed by an expert reviewer). Functional MRI assessments of language lateralisation address this issue by calculating laterality indices across a range of statistical thresholds, and creating a weighted mean (Wilke & Schmithorst, 2006). Using this method it is also possible to view the distribution of LI values over a range of statistical thresholds, which can inform the level of confidence with which lateralisation results can be interpreted. It is possible that the same sort of method may help identify critical activation within language fMRI results, although this has not yet been investigated. Currently, there is no consensus on the 'correct' threshold to use for localising fMRI activation in the pre-surgical setting, and it is likely that thresholds will always need to be adjusted and selected on an individual basis. In order for fMRI to extend its role in the pre-surgical decision-making process however - from assessing language lateralisation to language localisation - it is important to decipher at which threshold activation should be considered 'true'. Both false positives and false negatives can have serious consequences in the pre-surgical setting. At worst, false positives may prevent surgery from going ahead or restrict the area resected - influencing long term seizure outcome. In contrast, false negatives may result in resection of eloquent cortex, which could have devastating consequences for post-surgical cognitive function. The issue of statistical thresholding is demonstrated in the case of Patient A. At a threshold of $p<0.05$ (uncorrected) small clusters of activation were apparent along the superior temporal gyrus in proximity to the dorsal edge of the lesion. These clusters were not apparent at a threshold of $p<0.01$ (uncorrected). When stimulated with ICM however, tested clusters proved critical to language function. This information was used by the surgeon to guide the extent of the resection. Although fMRI did detect activation in

this critical region, activation in this area may have been disqualified as noise due to its absence at higher thresholds. More work is required to validate fMRI with ICM; which itself is a threshold-dependent technique. The currently small literature on this topic is reviewed in the introduction to this thesis (section **Error! Reference source not found.**), and highlights the need for systematic comparisons of fMRI and ICM using the same tasks and similar testing procedures. I have shown it is feasible to use the Panda Games for both methods; these tasks should enable systematic validation of fMRI with ICM in the future. A comparison of results across thresholds for both methods could be helpful in developing our understanding of how sensitive fMRI is for detecting language sites relative to ICM. In the longer term, validation with post-surgical outcome will be required, in order to quantify the validity and reliability of pre-surgical language mapping with fMRI.

7. CONCLUSIONS

This pilot study shows the Panda Games produce scans of the same quality as tasks currently validated for pre-surgical assessments of language (verb generation), but have the advantage of being suitable for young patients (aged < 7 years; Patient A) and patients with cognitive impairment (Patient F). They provide good sensitivity and specificity for localising language regions, and also provide new information regarding the localisation and lateralisation of basal temporal regions associated with semantic processing. Further work is required to enhance the potential of language fMRI in very young patients and patients with challenging behaviour (e.g. ADHD). However, the Panda Games provide some indication that specific design aspects may encourage compliance in these children, including; the use of colourful pictures within language tasks and use of child-friendly and cartoon-like introductory pack which generates some feeling of excitement around performing the tasks. In the longer term it is imperative that more studies strive to validate fMRI results with invasive intracranial mapping and long term post-surgical outcome. Only once this has been done can we reduce the need for invasive assessments in a population of children already undergoing significant medical treatment. Encouragingly, I have shown it is feasible to use the same task battery across methods, enabling validation in future studies.

PART 4: EXPERIMENTAL

Chapter 7: A systematic comparison of functional activation and functional connectivity methods for investigating language development in healthy children and children with focal epilepsy

1. ABSTRACT

Neuroimaging studies of language have recently experienced a paradigm shift, from investigating the topology of network activation to investigating network synchrony (functional connectivity). Functional connectivity has not yet been applied to investigate language development however. In this chapter I aimed to systematically compare measures of language network topology and synchrony, in terms of explaining variance associated with language development and impairment. Forty eight healthy children (4-12 years) and 21 children with focal epilepsy (5-12 years) performed a semantic decision task during fMRI. Activation in core language nodes was extracted and used for paired-ROI functional connectivity analysis. Nodal activation and single connections were compared between groups, and were entered into Principal Component Analysis to extract activated and synchronised sub-networks. Factor loadings from extracted networks were analysed in relation to age, epilepsy and language ability. Dorsal and ventral activated and synchronised networks were extracted from PCA. Activation in both networks decreased with age, suggesting increased efficiency. Activation in the ventral language network was reduced in children with epilepsy and explained 47% of variance in language ability. There was no evidence for a relationship between functional connectivity and age, epilepsy or language ability. I conclude that the typically developing language system is composed of dorsal and ventral networks. The ventral network appears crucial to typical language development; it is vulnerable to childhood onset focal epilepsy and predicts language ability. Understanding how the ventral system develops may therefore be crucial for making predictions regarding language outcome in patients being considered for neurosurgery.

2. AIMS

In the introduction to this thesis I highlighted the need for a more comprehensive understanding of how the language network develops, and how it may be affected by childhood onset epilepsy. This information is invaluable for pre-surgical planning. Over the past ten years there has been a shift in the fMRI literature from investigating network activation to network synchrony, using connectivity methods. However, there have been no systematic comparisons of functional activation and functional connectivity, to suggest which measure is most sensitive for investigating individual variability in language. In this chapter I aim to provide a systematic comparison of methods investigating the spatial and temporal organisation of language networks using conventional fMRI analyses and functional connectivity, respectively. By investigating the relation of these measures to a) age, b) epilepsy diagnosis and c) cognition, I aim to identify the optimal method for use in my own investigations of typical language development (Chapter 8). I also use this chapter as an opportunity to review the white matter architecture of the language network, which may contribute to post-surgical language outcome (although this has received relatively little attention in the literature so far).

3. INTRODUCTION

A comprehensive review of functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) studies of language over the last twenty years concluded; "*The next 20 years will need to focus on understanding how different brain regions interact with one another, and how specialisation for language arises at the level of distinct patterns of activation*" (Price, 2012). Functional connectivity (FC) is a measure of how tightly distinct regions of the brain are synchronised (Friston et al., 1993), providing a window into network dynamics. Initial studies suggest FC in the language network may change from childhood to adulthood in the typically developing population (Friederici, Brauer, & Lohmann, 2011). However, exactly how the language network changes – and how it responds to the effects of childhood onset epilepsy – are not yet known. In this chapter I aimed to investigate the spatial and temporal dynamics between classic language regions in frontal, parietal and temporal

cortices in the core (dominant) language network, and how this relates to age, epilepsy and language ability. To do this, I performed principle component analysis (PCA) on both fMRI activation and connectivity data to move beyond single region and single connection comparisons, and investigate organising principles in the developing language network. This also allowed for a systematic comparison of the relative importance of spatial and temporal network characteristics to understanding brain development, function and response to epilepsy. In doing so, I aimed to identify whether measures of network topology (activation) or synchrony (functional connectivity) provide the most valuable insight into language network development in healthy children and children with epilepsy. I reviewed the literature on structural networks supporting the mature language system to formulate hypotheses regarding functional networks supporting the language system. I also reviewed the current evidence for FC changes in adults with epilepsy, to formulate hypotheses for my sample of children with focal epilepsy. These are summarised below.

3.1 Structural connectivity of the language system

3.1.1 Connectivity of Broca's area

Broca's area, a classic language region in the left inferior frontal gyrus (IFG), can be parcellated anatomically - according to its cytoarchitecture - into Brodmann area's 44, 45 and 47 (Brodmann, 1909). Broca's area can also be parcellated according to function: There is an anterior-inferior to posterior-superior gradient within the IFG, such that BA 47, 45 and 44 are involved in semantic, syntactic and phonological processing, respectively (Bookheimer, 2002). More recently it has been shown that these cytoarchitectonically and functionally defined sub-regions within Broca's area are supported by distinct patterns of structural connectivity (Anwander, Tittgemeyer, von Cramon, Friederici, & Knosche, 2007; Brauer, Anwander, & Friederici, 2011; Frey, Campbell, Pike, & Petrides, 2008; Saur et al., 2008; Saur et al., 2010; Weiller et al., 2009). The most inferior and anterior portion of classic Broca's area (BA 47) is connected to the temporal lobe (TL) and inferior parietal lobe (IPL) by a ventral pathway system, comprised of several potential tracts; the inferior frontal occipital fasciculus (IFOF), extreme

capsule (EmC), middle longitudinal fasciculus (MLF) and inferior longitudinal fasciculus (ILF) (Catani & Mesulam, 2008; Friederici, 2009; Makris & Pandya, 2009; Makris et al., 2007, 2009; Schmahmann et al., 2007). The uncinate fasciculus (UF) – the last white matter tract to mature in the brain (Ellmore et al., 2010) - also connects BA 47 to the anterior temporal lobe. As its projections run medially to the hippocampus and amygdala, it is not considered to be part of the ventral language pathway by all (Saur et al., 2008). However, the UF may well contribute to language performance through its role in verbal memory (McDonald et al., 2008). The more posterior portion of classic Broca's area (BA 44) is connected to the IPL and posterior TL via a dorsal language pathway, comprised of the three superior longitudinal fasciculi (SLF I, II and III) and the arcuate fasciculus system (AF), which includes an anterior segment (directly connecting BA 44 to the posterior TL) and a posterior segment (indirectly connecting BA 44 to the posterior TL via the IPL) (Catani, Jones, & ffytche, 2005; Friederici, 2009; Makris & Pandya, 2009; Makris et al., 2005).

3.1.2 Connectivity of Wernicke's area

A functional gradient has also been shown in the temporal lobe (TL), such that the anterior TL is involved in semantic processing and the posterior TL in phonological processing (Scott, Blank, Rosen, & Wise, 2000; Wise et al., 2001). This anterior-posterior distinction has been further supported by cortico-stimulation studies (Corina et al., 2010). These findings suggest, like Broca's area, classic Wernicke's area may be divided into functional sub-regions. Whether these sub-regions are supported by distinct patterns of structural and functional connectivity, as in Broca's area, remains to be seen however. Studies investigating the structural and functional connectivity of Broca's area (discussed above) suggest the anterior temporal lobe (aTL) will show strong functional connectivity with BA 47/45, due to anatomical connectivity via ventral white matter tracts, and based on findings of co-activation of these regions during semantic processing. The posterior TL (pTL) in contrast is expected to show tight coupling with posterior regions involved in sound-to-motor mapping (the IPL and BA 44) supported by the dorsal SLF/AF systems. The middle TL (mTL) is extremely well connected, both anatomically and functionally, with other

temporal, frontal and parietal regions in adults (Turken & Dronkers, 2011). Further, patient case studies and intracranial mapping suggests this region is crucial to language function (Corina et al., 2010; Dronkers, Wilkins, Van Valin Jr., Redfern, & Jaeger, 2004; Geschwind, 1970; Wernicke, 1874). As such, the mTL may represent a critical 'hub' in the language network, as the cortical bridge between dorsal and ventral white matter tracts, creating a loop system (Weiller, Bormann, Saur, Musso, & Rijntjes, 2011). There are no studies, to my knowledge, investigating how the mTL becomes established as a critical network node during language development. This is important, as it may contribute to our understanding of language reorganisation in surgical candidates.

3.1.3 Connectivity of Geschwind's territory

The inferior parietal lobe, or more specifically the supramarginal gyrus (SMG) and angular gyrus (AG), form Geschwind's territory (Catani et al., 2005). Situated at the junction of visual, somatosensory, motor and auditory processing regions, this "*association area of association areas*" permits the cross-modal integration of information Geschwind believed key to human language development (Geschwind, 1965, page 641). A review of 1135 activation foci from 187 fMRI and positron emission tomography (PET) studies of semantic processing found a strong clustering of foci in the left SMG and AG, suggesting the IPL does indeed form a crucial part of the semantic system (Binder, Desai, Graves, & Conant, 2009). Based on a review of neurological and functional imaging studies, Binder and colleagues conclude the AG supports concept retrieval and integration, at the top of the processing hierarchy. As such it is proposed to play a role in behaviours requiring fluent conceptual combination, such as sentence comprehension, discourse, problem solving, and planning. In contrast, Binder and colleagues propose a more specific role for the SMG in semantic processing. Due to its anatomical proximity to somatosensory association cortices, they suggest the SMG stores abstract somatosensory knowledge acquired during performance of complex motor sequences, such as articulation. As such, the SMG is thought to play a role in a perception-action cycle (Fuster, 1998; Fuster & Oxford University, 2003), underlying phonological working memory and decision strategies (Price, 2012; Vigneau et al., 2006). Both

the SMG and AG are anatomically connected with BA 44 via the SLF/AF (Catani et al., 2005), despite their differing roles in phonological and semantic processing; supporting the concept of multiple dorsal sub-networks (Vigneau et al., 2006). The IPL also demonstrates ventral structural and functional connectivity, with the fusiform gyrus in the temporal lobe (Wang et al., 2012); another region implicated in Binder's model of the semantic system (Binder et al., 2009). However, research on connectivity of the IPL is currently scarce.

3.1.4 Organising principles of the language network: A multiple systems approach

Functional neuroanatomical models of language propose a division of labour in the mature language network between the ventral and dorsal systems, which are supported by the dorsal and ventral white matter tracts described above (Hickok & Poeppel, May, 2007). The dorsal/ventral distinction is not a simple one however: The dorsal and ventral language systems are themselves formed of multiple sub-routes, serving various language functions, as discussed previously (Price, 2012; Vigneau et al., 2006). The regional weighting of activation throughout these functional sub-networks appears to be modulated by task demands (Vigneau et al., 2006). Broadly however, the ventral system is proposed to support semantic processing (sound to meaning mapping) while the dorsal system has been implicated in sound to motor mapping (Saur et al., 2008), phonological and semantic working memory (Vigneau et al., 2006), hierarchical processing and processing of complex syntax (Friederici, 2011). Although highly controlled language tasks can distinguish between dorsal and ventral systems (Saur et al., 2010), it is likely both support naturalistic use of receptive and expressive language, in a highly integrated dynamic system. The middle temporal gyrus (MTG) is a prospective substrate for integration of dorsal and ventral systems, due to its extensive anatomical connectivity, and may be a critical language hub (Turken & Dronkers, 2011). The IFG may also support unification of information across language sub-networks (Friederici, 2011; Hagoort, 2005).

3.2 Language network connectivity and epilepsy

Children with focal epilepsy have an increased risk of language impairment (Parkinson, 2002). They also show an increased prevalence of atypical language organisation compared to typically developing children with conventional fMRI (Rosenberger et al., 2009). However, it is not known how childhood-onset epilepsy affects FC within the language system. I therefore reviewed the small literature on language network FC in adults with epilepsy, to draw hypotheses for my paediatric sample.

3.2.1 Decreases in functional connectivity

The language network is less well synchronised in adults with epilepsy, even when regional activation on conventional fMRI appears normal (Pravatà et al., 2011; Vlooswijk et al., 2010; Waites et al., 2006). Connectivity of the left hemisphere language network appears most severely affected in all of these studies, regardless of side of epilepsy. Crucially, it has been shown that poor language function in adults with epilepsy correlates with reduced FC (synchrony) between left hemisphere language regions (Pravata et al., 2011, Vlooswijk et al., 2010).

3.2.2 Increases in connectivity

There is some evidence that people with epilepsy may show increases in FC which extend beyond the typical language network. For example, both the posterior temporal lobe (Pravatà et al., 2011) and IFG (Waites, Briellmann, Saling, Abbott, & Jackson, 2006) show increased FC with the posterior cingulate cortex (PCC) during resting state fMRI in adults with epilepsy. The PCC is typically associated with the default mode network, is highly active during rest, and is typically deactivated during task performance (Waites et al., 2006). As such, the authors interpret these connections as maladaptive (Pravatà et al., 2011).

3.3 Hypotheses

3.3.1 The typically developing language system

The novel application of principal component analyses (PCA) to network activation and connectivity data in this chapter will allow the identification of networks containing clusters of brain regions with shared variability in signal change or synchrony during performance of a language task. Using this exploratory approach I aim to investigate whether distinct dorsal and ventral systems are identifiable during development. I hypothesise networks extracted from PCA will reflect a division of labour between ventral (BA 45/47, aTL /mTL) and dorsal (BA 44, mTL/pTL and AG/SMG) processing streams in the developing language network, in line with current functional neuroanatomical models of language. Combined fMRI and diffusion weighted imaging suggests there is a developmental shift from the ventral system during childhood to the dorsal system during adulthood, due to the relatively late maturation of the SLF/AF versus the EmC (Brauer et al., 2011). As such, I also hypothesise a developmental shift such that younger children show increased activation and synchrony within a ventral system, and older children within a dorsal system. I hypothesise the mTL and IFG will be the most well-connected nodes, representing network hubs. Based on evidence for different connectivity profiles within classic language regions, I will use small regions-of-interest (ROIs) which divide classic Broca's and Wernicke's areas - as well as Geschwind's territory (Catani, Jones, & ffytche, 2005) - into functional sub-regions, guided by previous cytoarchitectonic, anatomical connectivity and functional imaging studies (Anwander, Tittgemeyer, von Cramon, Friederici, & Knosche, 2007; Bookheimer, 2002; Brodmann, 1909; Catani et al., 2005; A. S. Dick & Tremblay, 2012b; Frey, Campbell, Pike, & Petrides, 2008; Geschwind, 1965; Saur et al., 2008; Vigneau et al., 2006). These small ROIs are hypothesized to reflect the functional subcomponents of the language network, thus providing a more detailed estimation of functional networks supporting task performance.

3.3.2 The language system in children with epilepsy

Increases and decreases in FC found by the small number of studies investigating language network dynamics in adults with epilepsy, suggest FC is an important measure for understanding the brains response to neurological disease, including intra- and inter-hemispheric reorganisation and functional compensation. However, there have been no previous reports of the relationship between measures of FC and language in children with epilepsy. I aimed to build on previous work by investigating the relationship of language network FC to epilepsy and cognition in children, and when controlling for the effect of language reorganisation. I hypothesise altered patterns of FC relative to healthy children, including decreased synchrony among left hemisphere language regions, in relation to poorer cognitive function.

4. METHODS

4.1 Participants

I analysed data which had been collected prospectively as part of a study investigating language development in healthy children and children with epilepsy at the Children's National Medical Centre, Washington DC, under the supervision of Dr Berl and Dr Gaillard. Data were available for 48 healthy children (4-12 years) and 21 children with left focal epilepsy (5-12 years) (Table 26). All participants were native English speakers with negative MRI findings (at 1.5T), did not have comorbid mood disorders or any systemic disorder which might affect central nervous system function. All participants with epilepsy were deemed to have focal epilepsy with impaired consciousness (Berg et al., 2010) based on medical history, clinical examination and (video) EEG, with normal MRI and unknown cause. Lateralization and localization of epilepsy was based on clinical description of seizures, post ictal physical exam when available, and (video) EEG. The majority of patients (76%) experienced left hemisphere seizures. Seizure onset was most often localized to the frontal or temporal lobes (81% of all patients; 75% of patients with left hemisphere onset). Patients were taking 1-3 antiepileptic medications at the time

of scanning (mode = 1): one patient was on Topiramate. All patients were seizure-free in the 24 hours prior to fMRI scanning.

	Patients (N=21)	Controls (N=48)	*p*
Descriptive statistics			
Age *(years)*	8.67 (±2.2)	9.06 (±2.48)	*0.54*
Sex *(females, males)*	9,12	24,24	*0.72*
Language lateralisation (laterality indices)			
Broca's Area	0.68 (±0.21)	0.71 (±0.16)	*0.37*
Wernicke's Area	0.54 (±0.42)	0.63 (±0.32)	*0.42*
Neuropscyhological performance			
General verbal ability composite *(Z score*	-0.73 (±0.8)	0.32 (±0.9)	*<0.001*
VIQ *(SS)*	100 (±13)	117 (±16)	*<0.001*
CELF-IV Core language *(SS)*	100 (±12)	113 (±11)	*<0.001*
EOWPVT *(SS)*	98 (±14)	114 (±16)	*<0.001*
Performance (non-verbal) IQ *(SS)*	99 (±11)	110 (±14)	*0.002*
In-scanner measures			
Movement *(mm)*	1.87 (±2.88)	1.97 (±3.2)	*0.90*
Task performance *(% correct)*	65.56% (±19)	78.88% (±17)	*0.008*
Epilepsy characteristics			
Age of seizure onset *(years)*	5.01 (±2.85)		
Duration of seizures *(years)*	4.23 (±3.54)		
Seizure frequency			
Daily	5%		
Weekly	14%		
Monthly	33%		
≤1 per 6 months	43%		
Seizure lateralisation			
Left	76%		
Right	24%		
Seizure localisation			
Frontal	14.30%		
Temporal	57.10%		
Parietal	4.80%		
Fronto-temporal	9.50%		
Temporo-parietal	4.80%		
Central	4.80%		
Undetermined	4.80%		

Table 26: Descriptives statistics for a sample of healthy children and children with focal epilepsy. Standard scores (SS) are shown for neuropsychological assessments.

4.2 Language dominance

Only children with left hemisphere language dominance were included, to control for the influence of language reorganisation when investigating intra-hemispheric connectivity. Language lateralisation was determined using a threshold-independent laterality index as described in Chapter 1. Language dominance was categorised according to the laterality index calculated in the inferior frontal gyrus (Broca's area).

4.3 Neuropsychological testing

Intelligence and language were assessed according to standardized administration with age appropriate measures: IQ was assessed with either the Differential Scales of Ability (DAS, ages 4–5 years) or the Wechsler Abbreviated Scale of Intelligence (WASI, ages 6–12 years) (Elliott, 1990; Wechsler, 1999). Fundamental language skills were assessed with the Clinical Evaluation of Language Fundamentals Preschool Version (CELF-P, age 4 years) or Fourth Edition (CELF-IV, ages 5–12 years) (Semel, Wiig, & Secord, 2003a; Wiig, Secord, & Messing Semel, 2004). Single word retrieval (naming) was assessed using the Expressive One Word Picture Vocabulary Test (EOWPVT, ages 2-80 years) (Gardner, 1990). Verbal measures (verbal IQ, core language ability and word retrieval) were combined using PCA to produce a general verbal composite score, which accounted for 74.36 % of the variance. This composite score was used to reduce the number of comparisons made when investigating the relationship of network measures to cognition.

4.4 Functional MRI parameters

Functional data were acquired using a 3.0 Tesla Siemens Magnetom Trio equipped with a standard CP head coil. Anatomical images of participants were collected using a sagittal T1 MPRAGE sequence, slice thickness of 1.0 mm, TR of 1600 ms and TE of 3.37 ms, which served to screen for anatomical abnormalities. Blood oxygen level-dependent (BOLD) changes were measured using a whole brain EPI

sequence with parameters: TR = 3000 ms, TE = 30 ms, FoV = 192 mm, and effective voxel size = 3.0 x 3.0 x 3.0 mm. Axial images were collected parallel to the anterior commissure–posterior commissure plane, which served as an origin of reference. Whole brain volumes consisted of 50 axial slices of 2.8 mm thickness with a 0.2 mm gap between slices.

4.5 Functional MRI language task

All children had a mock scanning session to enhance compliance and task performance. Participants performed the Auditory Description Decision Task (ADDT), during which they heard pre-recorded sentences of true or false definitions of common objects (e.g. "A long yellow fruit is a banana"). Participants were instructed to press the button on the MRI compatible response box when definitions were true. This task engages semantic processing regions along the temporal and inferior parietal cortices associated with sentence comprehension, and inferior frontal and inferior parietal regions associated with semantic predictions and word retrieval (Berl et al., 2014; Gaillard et al., 2007). This task also offers the benefit of performance monitoring and developmentally matched difficulty levels: *easy* (for 4-6 year olds), *medium* (7-9 year olds) and *hard* (10-12 year olds). Seventy percent of items were correct targets and 30% were foils. During the baseline condition participants listened to reversed speech and pressed the button upon hearing a tone. This allowed for the subtraction of bilateral activation due to auditory processing and was matched to the experimental condition for motor response, length of utterance and volume of presentation. The task was presented in a blocked design with ten alternating 30s blocks, five active and five baseline, resulting in a total scan time of five minutes. Individual stimuli were presented every three seconds with a total of 10 stimuli per block. This task was performed as part of a larger neuroimaging protocol performed during the same scanning session.

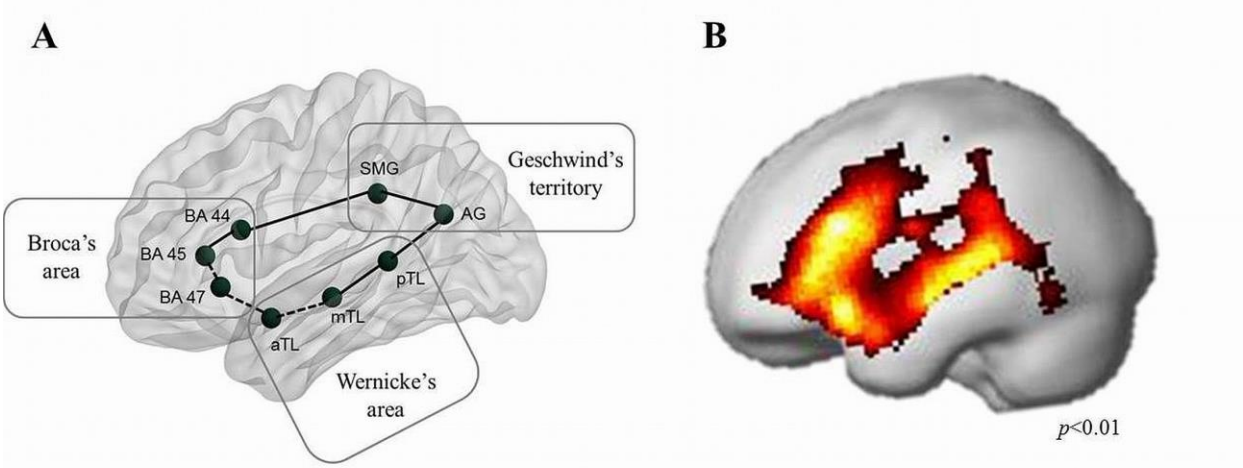
4.6 Functional MRI pre-processing

Image data pre-processing and group analyses were performed using Statistical Parametric Mapping software (SPM8; University College London) and Matlab (The MathWorks,Inc.: Natick, MA), as described in Chapter 3. Mean in-scanner

movement (as assessed by realignment parameters from spm8) did not differ between controls and patients (Table 26).

4.7 Left-Hemisphere regions-of-interest

I created eight anatomic ROIs in the left hemisphere (Figure 55 A) using the automatic anatomic labels (Tzourio-Mazoyer et al., 2002) in the Wake Forest Pick Atlas (Maldjian, Laurienti, & Burdette, 2004; Maldjian, Laurienti, Kraft, & Burdette, 2003), according to known anatomical boundaries (Brodmann, 1909) within Broca's area (inferior frontal gyrus), Wernicke's area (middle and superior temporal gyri) and Geschwind's territory (inferior parietal lobe). Anatomic ROIs were overlaid on the control group activation map (Figure 55 B) and a 4mm radius sphere was drawn around the global point maxima within each region. These ROIs were used as network nodes for the analysis of mean fMRI signal change and functional connectivity in the dominant (left hemisphere) language network. I classified nodes as dorsal or ventral (Figure 55 A) to interpret results in relation to current functional neuroanatomical models of language (Hickok & Poeppel, 2004). To ensure placement of ROIs was not unduly influenced by our sample, I confirmed these were common areas for language activation by entering peak coordinates into an fMRI database (http://beta.neurosynth.org) which verified >70 studies of language also found activation within 6 voxels of these activation peaks.



C

Node characterisitcs and function

Node	*Co-ordinates*	*Language function (from Price 2010; 2012)*
aTL	-56, 4,-16	Semantic associations
mTL	-54,-20,-8	Semantic associations and sentence-level integration
pTL	-56,-42,6	Semantic access (especially at sentence-level)
AG	-42,-64,24	Semantic integration (cross-modal and sentence-level), and top-down semantic prediction
SMG	-44,-38,32	Articulatory loop (verbal working memory), and subvocal articulation to aid comprehension
BA 44	-54,16,18	Integration (of inputs, expectations and meaning), and top-down semantic prediction
BA 45	-48,30,8	Semantic decisions
BA 47	-44,24,-4	Semantic (word or concept) selection/retrieval

Figure 55: Network nodes (ROIs) placed within classic language regions (A) according to peak maxima in the control group activation map (B), and their functional characteristics (C). ROIs were placed along the dorsal (solid line) and ventral (dashed line) language streams (A). Co-ordinates in MNI space and the associated roles of each node in the language network are listed (C), according to consistent structure-function relationships which have emerged from 20 years of language fMRI and PET studies (taken from Table 4 in Price 2012, and also Price 2010). Network nodes include; BA 47 (Pars Orbitalis), BA 45 (Pars Triangularis), BA 44 (Pars Opercularis), aTL (anterior temporal lobe), mTL (middle temporal lobe), pTL (posterior temporal lobe), SMG (supramarginal gyrus) and AG (angular gyrus).

4.8 Region-of-interest analyses

4.8.1 Nodal activation

Percent mean fMRI signal change was extracted from each ROI using MarsBaR software (Brett et al., 2002). I used repeated measures analysis of variance (ANOVA) to investigate the effect of group (between-subject factor) and node (within-subject factor) on activation strength.

4.8.2 Paired-ROI functional connectivity analysis

Mean time courses were extracted from each of the eight network nodes. Pearson's *r* correlation coefficients were calculated between mean time courses for all node pair combinations (25 in total) as a measure of functional connectivity strength, and were converted to Fisher's *z* scores for statistical analyses. Mixed ANOVA were performed to investigate the effect of group (between-subject factor), connection (within-subject factor) and connection type (within-subject factor; fronto-temporal, fronto-parietal and temporo-parietal) on functional connectivity strength.

4.8.3 Identification of language networks using Principal Component Analysis

Two principal component analyses (PCA) were performed across all participants (N=69), using a direct oblimin rotation with 25 iterations and corrected with the Kaiser criterion (including variables with an eigenvector >1). Only variables with factor loadings >0.4 (which explain at least 16% of the variance in a given factor) were reported (Field, 2005). *Activated networks* were investigated (PCA-A) by entering mean fMRI signal change values for each node. Factors from this analysis represent a cluster of brain regions which behave similarly across individuals (they share between-subject variability in task-related fMRI signal change). As it is their collective response which defines them as a single factor, we interpret these regions as representing a functional network. *Synchronised networks* were investigated (PCA-S) by entering all 25 *z* scores from the paired-ROI functional connectivity analysis. Here factors represent networks of connections with shared between-subject variance in functional connectivity

strength. Regression factor scores for individual participants were generated for each PCA (A and S), to quantify how strongly the extracted networks were activated (PCA-A) or synchronised (PCA-S) for each individual. A factor score of 0 represents the average for the sample. Positive factor scores indicate higher-than-average activation or synchrony, and negative factor scores indicate lower-than-average activation or synchrony.

4.9 The relation of language networks to age, epilepsy, performance and cognition

Two mixed ANOVA were performed to investigate the effects of component (network), diagnostic group and age group (4-6, 7-9 and 10-12 years) on regression factor scores from PCA-A and PCA-S. The association of age of onset, duration of epilepsy, in-scanner task performance and general verbal ability to factor scores was investigated using Pearson's r correlation coefficient. Networks showing a significant relationship with the general verbal ability composite were investigated further in relation to individual neuropsychological measures. This stepwise approach served to reduce the number of comparisons made. Finally, all networks were investigated as predictors of language function (Core Language composite scores on CELF-IV) with multiple linear regression, to identify which network and which parameter (fMRI signal change or functional connectivity) were most informative for predicting language outcome in children with epilepsy. All *p* values were adjusted for multiple comparisons using the Bonferroni procedure.

5. RESULTS

5.1 Neuropsychological performance

Patients had lower verbal and nonverbal abilities than controls (Table 26). Verbal ability (as measured by the composite score) did not differ between age groups and did not correlate with age in controls (p=0.76 and 0.17, respectively) or patients (p=0.75 and 0.81, respectively). Verbal ability was not correlated with age of onset or duration of epilepsy in patients (p values >0.42). Task performance did not differ between age groups in controls (accuracy p=0.08, reaction time p=0.48) or patients

(accuracy p=0.32, reaction time p= 0.44). In the control group, there was a trend for a correlation of performance accuracy with age (p=0.06) which was driven by one outlier. There was no correlation once this outlier was removed (p=0.21).

5.2 Group Activation

Both groups activated the language network during ADDT performance, including left superior temporal regions, temporal pole, inferior parietal lobe, Broca's area (left BA 44 and 45) and middle frontal gyrus, bilateral BA 47, basal ganglia, anterior cingulate cortex and right cerebellum (Figure 55 B).

5.3 Region-of-interest analyses

5.3.1 Nodal activation

Patients showed lower mean fMRI signal change than controls (main effect of group: $F(1,64) = 10.74$, p=0.002). Post-hoc independent samples t-tests confirmed signal change values were lower in patients for all nodes (p values <0.02) except BA44 (which was higher in patients; p=0.047) and the angular gyrus (which did not differ between groups; p=0.95) (Figure 56 A).

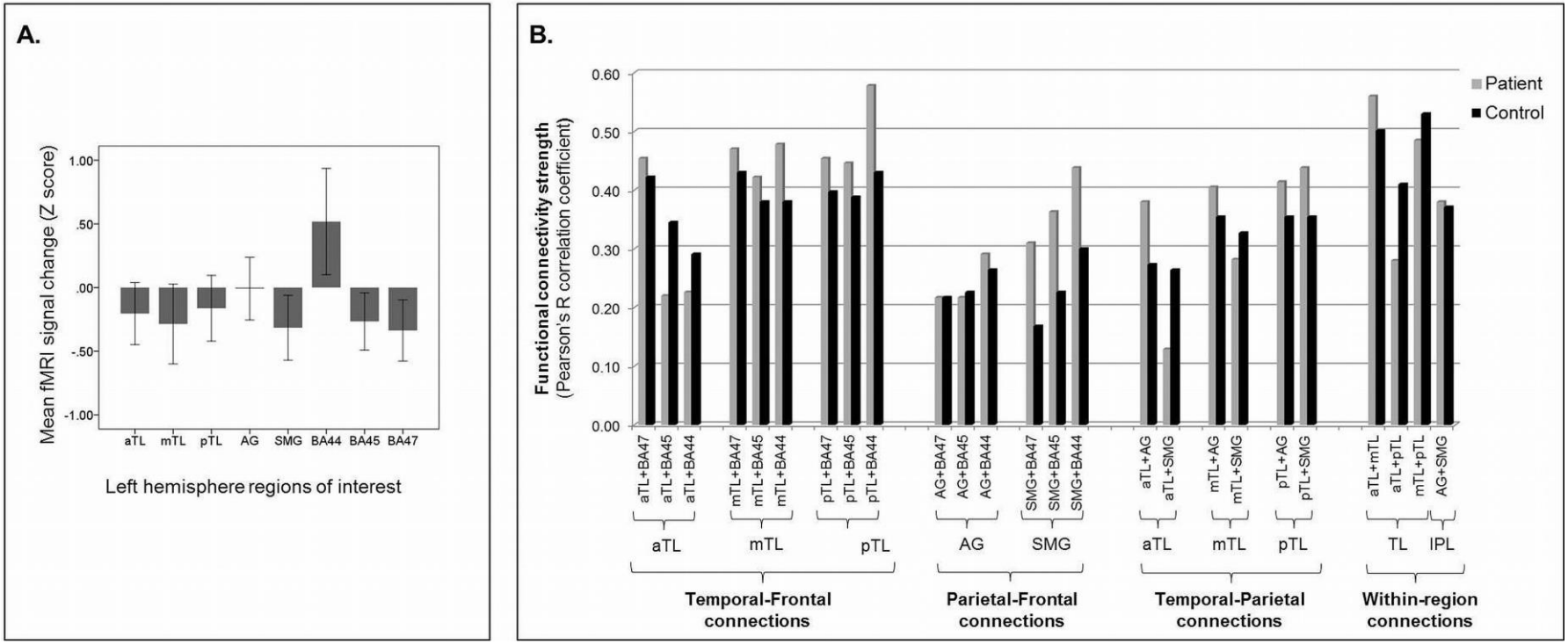


Figure 56: Mean fMRI signal change (Z scores) in network nodes in the patient group. Z scores show the mean fMRI signal strength in network nodes for patients, expressed in relation to the control group mean and standard deviation (with 95% confidence intervals). B) Strength of individual functional connections by region and by diagnostic group. Pearson's R correlation coefficient values for patients (grey) and controls (black); no significant group differences.

5.3.2 Paired-ROI functional connectivity analysis

Mean fMRI signal time courses were positively correlated for all 25 network node pairings (*p* values <0.001). There was no main effect or interaction effect of diagnostic group on overall functional connectivity ($F(1,67) = 0.34$, p=0.56), nor on fronto-temporal, fronto-parietal or temporo-parietal connections (all *p* values > 0.31) (Figure 56 B).

5.3.3 Identification of language networks with PCA

Two factors were extracted from both PCA-A and PCA-S (Figure 57): dorsal and ventral activated networks (dA and vA, respectively) and dorsal and ventral synchronised networks (dS and vS, respectively). To confirm this factor solution provided a valid representation of the typical network structure, we re-calculated PCA-A in healthy controls alone. Encouragingly, this analysis also extracted two factors (explaining 55% of the variance) which distinguished between dorsal and ventral regions: The dorsal factor showed activation in the inferior frontal gyrus (BA 44 and 47) and inferior parietal lobe (AG and SMG) and explained 39% of between-subject variability. The ventral factor showed combined activation of all temporal nodes and BA 45, and explained 16% of between-subject variability.

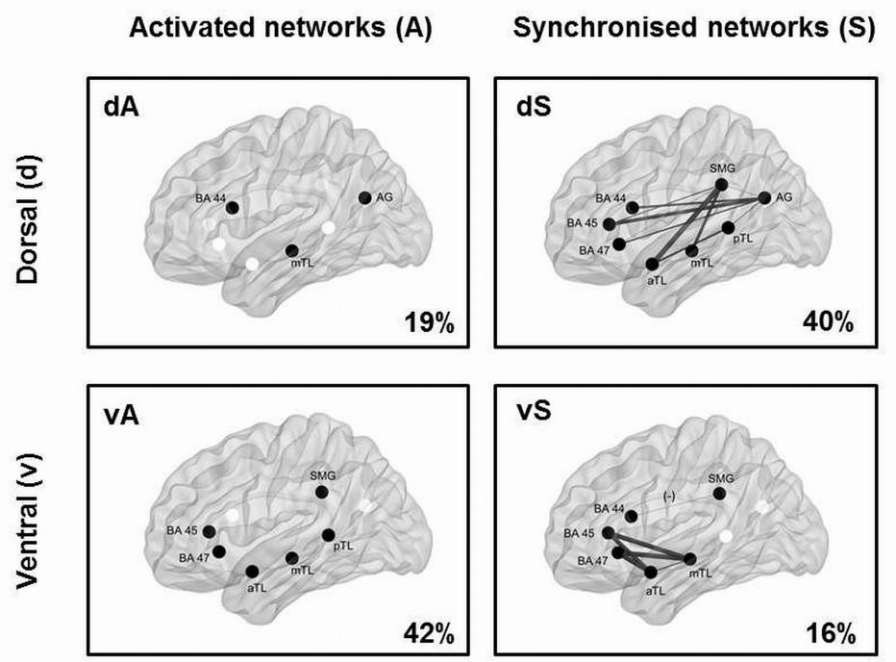


Figure 57: Language networks extracted by Principal Component Analyses. Dorsal (d; top row) and ventral (v; bottom row) principal components extracted from PCA-A (left) and PCA-S (right). Components from PCA-A (activated networks; A) represent networks of nodes (black) with shared between-subject variability in task-related fMRI signal fluctuations across individuals. Components from PCA-S (synchronised networks; S) represent networks of node-pairings with shared between-subject variability in functional connectivity. The percentage of between-subject variability explained is depicted for each component (bottom right corner). For synchronised networks, line thickness is weighted by mean factor loading, where thicker lines indicate higher factor loadings. Negative factor scores are highlighted (-) for vS.

5.4 Activated language networks

5.4.1 Effect of age

An interaction of age and diagnostic group ($F(2,60)=3.74$, $p=0.03$) showed that, in the control group, younger children (4-6 year olds) had higher factor scores than older children (10-12 year olds) ($p=0.02$; Figure 58 A). Correlation analyses also

showed a decrease in factor loadings with age in the control group, for both components (r=-.38 and -.34, *p* values < 0.03; Figure 58 B). This correlation remained significant when controlling for task performance (accuracy and reaction time) and verbal ability (as measured by the composite score) for vA (r=-.32, *p*=0.045) and at trend level for dA (r=-.27, *p*=0.09). There were no age effects in the patient group.

5.4.2 Effect of epilepsy

There was an interaction of diagnostic group and component ($F(1,60)=47.1$, $p<0.001$; Figure 58 A), suggesting between-group differences in factor scores for dorsal and ventral networks. Post-hoc ANOVAs investigating the effects of age and epilepsy on factor scores for each component separately revealed children with epilepsy (regardless of age) had greatly reduced factor scores for vA compared to healthy children ($F(1,60)=44.3$, $p<0.001$; Figure 58 A, bottom row), even when controlling for group differences in task performance ($p<0.001$). In contrast, for dA there was a trend for an interaction of age and group, suggesting older children with epilepsy (> 7 years of age) showed increased activation in the dorsal network relative to healthy children of the same age ($F(2,60)=2.69$, $p = 0.08$, Figure 58 A, top row). However, this was no longer significant when controlling for group differences in task performance accuracy ($p=0.14$). Further analyses showed factor scores for dA were strongly correlated with task performance accuracy for children with epilepsy (r=.62, *p*=0.008), but not for controls (r=-.08, *p*=0.63), such that children with epilepsy who performed better inside the scanner showed increased activation in this dorsal network. Neither vA nor dA were related to age of onset or duration of epilepsy.

5.5 Synchronised language networks

There was no main effect or interaction of component, age or diagnostic group on factor scores from PCA-S. Factor loadings from PCA-A and PCA-S were not correlated (r<.08, *p*>0.49) suggesting they measure distinct attributes of the language network.

5.6 The relation of network measures to task performance and language ability

Due to the main effect of diagnostic group on factor scores from PCA-A, correlation analyses were performed in control and patient groups separately. In-scanner task performance was not associated with network measures when controlling for age. Higher factor scores for vA were associated with better general verbal ability for controls (r=.30, p=0.04) and patients (r=.47, p=0.04). Specifically, higher VIQ (r=.29, p<0.05) and CELF-IV Core Language composite scores in controls (r=.31, p=0.04; Figure 58 C), and higher CELF-IV Core Language composite scores in patients (r=.69, p=0.002; Figure 58 C). The correlation of vA factor scores and Core Language composite scores remained significant when controlling for in-scanner task performance (controls: r=.35, p=0.02, patients: r=.75, p=0.002) and when controlling for age for patients (controls: r=.28, p=0.06, patients: r=.68, p=0.001). There were no significant correlations between factor scores from PCA-S and the general verbal ability composite.

5.7 Predictors of language ability

To investigate the independent contribution of different network measures (vA, dA, vS, dS) to language ability in children with and without epilepsy, stepwise linear regression was performed in both groups separately. In both groups, only activation strength in vA predicted language ability (CELF-IV core language composite scores); explaining 47% of language performance in children with epilepsy ($R^2_{adjusted}$ = 0.47, p=0.002) and 8% in healthy children ($R^2_{adjusted}$ = 0.077, p=0.04).

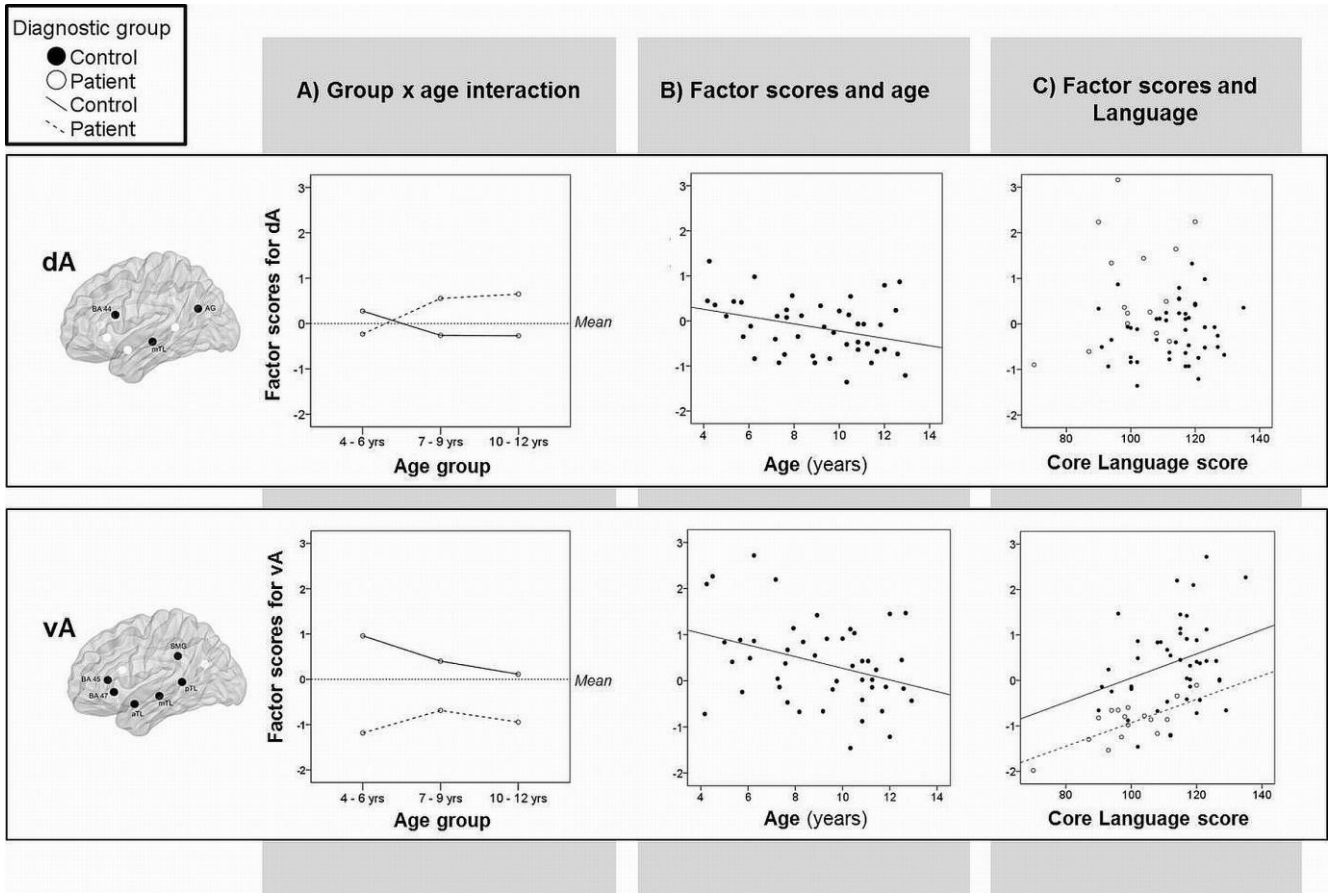


Figure 58: A) Significant interaction of epilepsy and age group on factor scores from dorsal (top) and ventral (bottom) networks extracted from PCA-A. Mean factor scores for controls (solid line) and patients (dashed line) on dA and vA are depicted across three age categories. Factor scores for vA were significantly reduced in the patient group (bottom), while patients showed a trend for higher factor scores on dA (top). In the control group, factor scores were significantly higher for 4-6 years old than 10-12 year olds (p=0.02) for both vA and dA. B) Significant correlation between factors scores from PCA-A and age in healthy children. C) Relationship of factor scores with language ability as measured by Core Language composite scores on the CELF-IV for dA (insignificant; p>0.05) and vA (controls: r=.31, p=0.04, patients: r=.69, p=0.002). The corresponding network is depicted next to the Y axis for each graph, to aid interpretation.

6. DISCUSSION

Using novel methods I provide new evidence for activated and synchronised networks, which suggest the developing language system - like the mature language system – shows a division of labour between dorsal and ventral systems (Hickok & Poeppel, 2004; Scott et al., 2000; Wise, 2003). While both networks were recruited to support language performance in typically developing children, particularly early in life (4-6 years), children with focal epilepsy failed to recruit the ventral system, and this was a marker of poorer language ability. These findings highlight the importance of the ventral language system for typical language development, and provide a potential explanation for the increased incidence of language impairment in children with epilepsy. Crucially, measures of network topology appear to provide vital information regarding how the language system matures, and how it responds to epilepsy.

6.1 Dorsal and ventral systems in the typically developing language system

Through the novel application of PCA to both fMRI signal change and functional connectivity data, I was able to extend findings from single-region and single-connection comparisons, to investigate developmental changes in language network topology and synchrony. Both PCA extracted two main components (networks), which I interpreted as having a dorsal or ventral stream topology, based on known white matter connections. These provide novel evidence for a degree of functional separation within the language system, previously only seen in structural connectivity studies. As language demands were not overtly manipulated, the function of each network cannot be inferred directly. However, structure-to-function mappings from the conventional fMRI literature (Price, 2012) provide a strong basis for interpreting the functional roles of these sub-networks.

6.1.1 Ventral network

I identified active (vA) and tightly synchronised (vS) networks predominantly involving BA 45 and 47, and the anterior and middle temporal nodes. These regions are implicated in semantic processing of sentences and semantic decision making, which are necessary skills for auditory comprehension. Component vA

also included activation in the posterior temporal and supramarginal nodes, implicated in sentence-level semantic access and sub-vocal articulation to support comprehension, respectively. I found a strong correlation between activation strength in vA and the Core Language Composite on the CELF-IV, which is comprised of several tasks requiring receptive language (following directions, recalling sentences etc.) and further supports the role of the ventral system in auditory comprehension. The supramarginal gyrus and BA44 are together implicated in a dorsal phonological working memory system (Price, 2012; Vigneau et al., 2006). Synchrony between these regions was attenuated in vS, which could reflect reduced demands on dorsal working memory networks or attentional systems when comprehension is achieved mainly via semantic processing in the ventral system.

6.1.2 Dorsal network

I identified a more dorsal network with activation increases in BA44, the angular gyrus and middle temporal node (dA). I also identified a widely connected network (dS) showing tight synchrony of inferior parietal nodes with inferior frontal and temporal nodes. This pattern of functional connectivity mirrors the underlying white matter tracts of the dorsal system (arcuate fasciculus), which connects the inferior parietal lobe to both the frontal and temporal lobes (Catani et al., 2005). As such, I consider this a dorsal network, which may be involved in working memory and phonological processing (Saur et al., 2008; Vigneau et al., 2006). Crucially, this network shows functional connectivity among both dorsal and ventral nodes, emphasising the overlap of these systems during naturalistic use of language (Weiller et al., 2009).

6.1.3 A middle temporal hub

Over the course of development the internal sensori-motor, memory and affective representations of words become strongly associated over time and due to extensive experience with language. Although highly controlled language tasks can distinguish between dorsal and ventral systems in healthy adults (Saur et al., 2010), these associations firmly integrate the two pathways into a synergistic system during naturalistic use of language (Pulvermüller & Fadiga, 2010;

Rolheiser et al., 2011). The middle temporal gyrus is a potential substrate for integration of dorsal and ventral systems due to its extensive anatomical connectivity, and may be a critical language hub (Turken & Dronkers, 2011). My findings support this hypothesis; the middle temporal node was involved in both dorsal and ventral networks identified with PCA.

6.2 Typical maturation of dorsal and ventral networks

6.2.1 Decreased whole-network activation with age

Longitudinal and cross-sectional studies in children using conventional fMRI have shown both increases and decreases in language network activation with age (Ahmad et al., 2003; Berl et al., 2010; Berl et al., 2014; Brauer & Friederici, 2007; Brown et al., 2005; Lidzba et al., 2011; Plante, Schmithorst, Holland, & Byars, 2006; Schmithorst et al., 2006; Szaflarski, Altaye, et al., 2012; Szaflarski, Schmithorst, et al., 2006; Yeatman et al., 2010). Findings are inconsistent and conflicting however, most likely due to use of different language tasks and inconsistent control over in-scanner task performance and language proficiency. For example, up to 50% of apparent age-related changes in network activation can be due to task performance differences (Brown et al., 2005). I therefore provide novel evidence that whole-network activation in dorsal and ventral language systems reduces with age during typical development, when task performance and language proficiency are well-controlled. Increased network activation in younger children may suggest the task was more difficult for these children (Yeatman et al., 2010), despite an age-adjusted task design and equivalent performance accuracy across age groups. Alternatively, reduced whole-network activation in older children may reflect refining of systems with age, with only a subset of nodes supporting proficient task performance later in life. For example, young children (<10 years) show activation along the extent of the inferior frontal gyrus during a syntactic processing task, whereas older children (> 10 years) show localised activation around BA 44 (Nunez et al., 2011). As such, my findings reported here could be considered as evidence supporting the interactive specialisation account of cognitive development, which predicts refining of cognitive networks with age (Johnson, 2011).

6.2.2 Increased whole-network activation with increased language proficiency

I also found evidence to suggest healthy children with more proficient language – regardless of age and in-scanner task performance – show increased activation throughout the ventral network. Conventional fMRI studies have also shown increased language proficiency is associated with increased and more extensive activity in left hemisphere language regions (Lidzba et al., 2011; Nunez et al., 2011; Yeatman et al., 2010).

6.2.3 No maturational change in functional connectivity

In contrast to findings of network activation, I did not find evidence for age-related changes in functional connectivity. Although left hemisphere functional connectivity differs when comparing very young children to adults and when using different methods (Friederici et al., 2011), another study which investigated changes across childhood using similar methods also failed to find evidence for major age-related connectivity changes (Wilke, Lidzba, & Krageloh-Mann, 2009). This may reflect the early establishment of fundamental white matter structures in the brain, despite their on-going refinement after 5 years of age (Lebel, Walker, Leemans, Phillips, & Beaulieu, 2008). Alternatively, the poor temporal resolution of fMRI or the influence of hidden dynamics may limit our ability to detect changes in network synchrony. Dynamic causal modelling (Friston, Harrison, & Penny, 2003) may be better placed to investigate subtle, or more complex, changes in network dynamics across childhood and adolescence, for example by accounting for possible age-related alterations in the haemodynamic response function (Friston, Moran, & Seth, 2013).

6.3 The effects of childhood-onset focal epilepsy on language network topology

6.3.1 Decreased activation of a ventral network

Principal component analysis suggested ventral regions of the language network work together in a co-ordinated network, which is compromised in children with epilepsy. The typical developmental trajectory of activation in this ventral

network for healthy children was not seen in children with epilepsy, nor was there evidence for developmental delay. Instead, activation of this network was consistently decreased for children with epilepsy across childhood, suggesting a pattern of atypical development. In support of this interpretation, reduced activation in this network was associated with poorer language function, predicting almost half the variance in language ability in children growing up with epilepsy. Previous studies also suggest the ventral system may be particularly crucial for typical language development, due to its relatively early maturation compared to the dorsal system (Brauer, Anwander, Perani, & Friederici, 2013). Further, the relevance of the ventral system to (receptive) language outcome has been documented in several patient populations (Agosta et al., 2013; Crinion & Price, 2005; Warren, Crinion, Lambon Ralph, & Wise, 2009). For example, in a large sample of adults with acute stroke, auditory comprehension deficits were associated with lesions affecting the extreme capsule (a ventral tract) (Kummerer et al., 2013). There is also accumulating evidence to suggest early structural damage to the ventral system has important consequences for language development. For example, damage to the extreme capsule and uncinate fasciculus - and not the arcuate fasciculus - is associated with language impairment in children born preterm (Northam et al., 2012). In addition, lesions encroaching on the ventral system are associated with functional reorganisation of language in children with medically resistant epilepsy (Pahs et al., 2013), suggesting damage to this system impacts the wider language network. Unlike these previous studies, our sample had no overt structural lesion. I therefore hypothesise changes in the ventral system in this study are related to epilepsy variables or its remote cause. Given the relatively short duration of epilepsy however I could not fully address the contribution of longer duration of epilepsy, which is best addressed by longitudinal studies. My initial findings however, do suggest the topology of ventral network activation may provide insights into the neural mechanisms of language impairment in this population.

6.3.2 Increased dorsal network activation

In contrast to ventral network nodes, my analysis of mean fMRI signal change showed activation in BA 44 was increased in patients. This region has been

associated with the dorsal stream, and is involved in processing complex syntax, syntactic working memory, phonological processing and lexical retrieval (Fiebach, Schlesewsky, Lohmann, von Cramon, & Friederici, 2005; Friederici, 2006; Friederici, Bahlmann, Heim, Schubotz, & Anwander, 2006; Heim, Eickhoff, & Amunts, 2009; Heim, Eickhoff, Friederici, & Amunts, 2009). Principal component analyses revealed BA 44 shared between-subject variance in task-related activation with the middle temporal node and supramarginal gyrus (dA), and was functionally connected (dS) with regions anatomically connected via the arcuate fasciculus. Activation of this dorsal network was increased in patients (especially those >7 years of age), providing novel evidence that reduced activation in one stream may be accompanied by increased activation in the other. Increased activation of the dorsal network was associated with improved task performance in the patient group. This finding suggests flexibility in recruitment of the two streams enables functional compensation.

6.4 The effect of childhood-onset focal epilepsy on language network synchrony

Previous studies suggest functional connectivity in the left hemisphere language network is reduced in adults with epilepsy regardless of side of seizure onset, and that this is associated with poorer language function (Pravata et al., 2011; Vlooswijk et al., 2010; Waites, Briellmann, Saling, Abbott, & Jackson, 2006). Contrary to these findings - and my own hypotheses - we found no group differences in functional connectivity at either level of analysis; single-connection (paired-ROI functional connectivity analysis) or network (PCA). Our findings may differ from previous studies because of four principal methodological differences:

6.4.1 Size of region of interest

Previous studies performed functional connectivity analyses using large anatomical ROIs, comprised of functionally distinct brain regions. Large ROIs may conceal changes in connectivity associated with specific functional networks: For example, an ROI based on the entire inferior frontal gyrus will average time courses from BA 44, 45 and 47, which show a posterior-anterior gradient for phonological, syntactic and semantic processing, respectively (Bookheimer,

2002). For this reason, I placed small ROIs within functional subdivisions of classic language regions. Large ROIs may also result in overlap with other (non-linguistic) networks supporting task performance, such as attention or executive functioning. Indeed, previous studies using large ROIs report changes in functional connectivity most often involving the anterior and posterior cingulate cortices (Vlooswijk et al., 2010, Waites et al., 2006). These regions may be involved in attentional processes supporting task performance, rather than linguistic processing per se (Leech et al., 2011, Fu et al., 2002). I aimed to retain the focus of our analysis to core language processing regions, by placing small ROIs within specific language regions.

6.4.2 The effect of language reorganisation

In contrast to previous studies, I investigated functional connectivity in children who retained left hemisphere language dominance. Because atypical (right or bilateral) language lateralisation is more common in people with epilepsy compared to healthy controls (Rosenberger et al., 2009), it is difficult to differentiate the effects of language reorganisation from the effects of epilepsy on functional connectivity in previous studies. The one study that did control for language dominance (i.e. only participants with left hemisphere language dominance were included) found no evidence for reduced left hemisphere functional connectivity in adults with epilepsy (Protzner & McAndrews, 2011). Previous findings of reduced functional connectivity between language regions in adults with epilepsy may therefore reflect differences in language dominance, rather than the effect of epilepsy per se. To further emphasise this point, I failed to find any association between epilepsy variables and network measures in this study. Even in the aforementioned studies of adults with epilepsy, neither age of onset, frequency of seizures, drug load, nor seizure focus accounted for variation in language network functional connectivity (Vlooswijk et al., 2010). However, initial research does suggest seizure duration (chronicity) may be an important variable for network connectivity; a prolonged duration of epilepsy has been shown to reduce functional connectivity (Luo et al., 2011) and promote a more random network configuration (van Dellen, 2009). In addition, a prolonged

duration of epilepsy may also be accumulatively detrimental to language function (Caplan, 2009). As such, it is possible that children in this study may not have experienced a sufficient duration of epilepsy to demonstrate changes in functional connectivity compared to adults with epilepsy. However, previous research in adults with epilepsy also provides little support for a direct influence of epilepsy on functional connectivity in the language system (discussed above).

6.4.3 Task performance difficulty and effort

In this study, participants performed an fMRI task which was adjusted according to developmental ability, to reduce variability of in-scanner task performance and effort. In this context, I was unable to identify group differences in functional connectivity, despite differences in network topology which remained highly significant when controlling for in-scanner performance. Studies which have reported reduced functional connectivity between core language regions in adults with epilepsy (Pravata et al., 2011; Vlooswijk et al., 2010) used a one-size-fits-all approach to task design, and did not control for in-scanner task performance. As such, differences in functional connectivity reported in these studies may reflect differences in performance accuracy or effort between healthy participants and participants with epilepsy. Indeed, changes in functional connectivity are related to poorer performance on similar language tasks outside the scanner (Vlooswijk et al., 2010).

6.4.4 The effect of cognitive impairment

In adults, changes in language network connectivity are more pronounced in patients with a greater degree of cognitive impairment (Vlooswijk et al., 2011). It is therefore possible that functional connectivity changes may be apparent in children with more severe epilepsy and cognitive impairment; for example, those with drug resistant epilepsy.

Aside from these methodological differences, my results suggest functional connectivity is less vulnerable to the effects of childhood onset epilepsy than adult onset epilepsy. This requires empirical support however, particularly with respect to the effect of epilepsy duration on language networks. Future studies investigating

the relationship of functional connectivity to disease status, task performance and language ability in adults compared to children would be highly valuable, to assess the sensitivity of connectivity measures to variation in disease state and ability across the lifespan.

7. CONCLUSIONS

Findings from this chapter highlight the value of fMRI measures of network topology (regional activation) for understanding development of language networks and their response to childhood onset epilepsy. Specifically, I found that failure to activate the ventral language system predicted poorer language ability in children with epilepsy. Activation strength in the ventral language system therefore represents an important target for the remaining chapters in this thesis, which aim to investigate developmental trajectories in the language system most relevant to children with epilepsy (Chapter 8). The ventral network can be investigated using the Listen and Name game, which investigates the effect of speech intelligibility on network activation (Scott et al., 2000; Scott, Rosen, Lang, & Wise, 2006; Vagharchakian, Vagharchakian, Dehaene Lambertz, Pallier, & Dehaene, 2012).

Chapter 8: Age-related changes in the auditory comprehension network, controlling for task performance difficulty and accuracy.

1. ABSTRACT

How structure-function mappings develop in the language system is not yet known, but is important for making predictions in the presurgical setting. In this chapter I aimed to investigate age-related changes in the typically developing ventral language system, while controlling for performance accuracy and effort. Thirty six healthy children (aged 5-16 years) performed the age-appropriate version of an auditory sentence comprehension task from the Panda Games battery (Listen and Name) during fMRI. Age related changes in network activation were investigated using multiple linear regression in spm8. Functional MRI showed a fronto-temporal comprehension network. Task performance accuracy was well controlled across age. Age-related increases of activation were seen associated with overall task performance (including sensori-motor and executive aspects of the task) in frontal regions (Pars Orbitularis, medial superior frontal gyrus and inferior frontal sulcus/precentral gyrus, SMA and cingulate cortex) and temporal regions (superior temporal gyri/sulci). Age-related increases in activation associated specifically with semantic processing were seen in bilateral mid-fusiform gyri and left posterior inferior temporal gyrus. Lateralisation of activation was stable across age. I conclude that late maturing brain regions associated with executive functioning and sensori-motor processing become increasingly active during auditory comprehension with age. Amodal semantic processing regions in the basal temporal cortex also become increasingly active, potentially reflecting increased cross-modal integration of semantic information with age. At most, age contributes only minimally to individual differences in lateralisation of the auditory comprehension network.

2. AIMS

This chapter aims to define age-related changes in the localisation and lateralisation of activation in the auditory comprehension network – and specifically in core semantic and syntactic processing regions – whilst providing stringent control over task performance difficulty and accuracy. In doing so, I aim to chart the typical trajectory for receptive language network development. Results are interpreted in the context of different frameworks for understanding cognitive development, and are discussed in terms of their relevance to pre-surgical decision-making.

3. INTRODUCTION

In order to define age-related changes in the auditory comprehension network, I reviewed fMRI studies of sentence-level auditory comprehension in healthy children and adolescents (aged 4-25 years). Papers which provide a summary or which aggregate analyses from previously published data are not reported (e.g. Holland et al., 2007). Word-level comprehension tasks were excluded. In total, I could find only 10 previously published studies investigating the development of sentence-level auditory comprehension (Table 27). The majority of these studies have been published by groups working on the interface of experimental research and clinical application, such that investigations of language in children are ultimately fuelled to identify markers of language impairment in patients or to inform the neurosurgical decision-making process. The motivations behind this research have shaped our current understanding of language network development in two ways. First, fMRI tasks have often been designed and optimised for clinical application and exploratory analysis, rather than testing specific developmental hypotheses. Results therefore tend to be over-inclusive, including activation associated with several levels of processing (not always limited to language). Consequently, although brain regions involved in auditory comprehension have been relatively well defined in children and adolescents, it has not been possible to chart the development of specific structure-function mappings within this system. Second, in children the principle clinical application of language fMRI is for assessment of language dominance in pre-surgical candidates (as a non-invasive alternative to the WADA test). Thus, there has been a large focus on age-related changes in lateralisation of function, rather than localisation of function. Below I summarise what is currently

known about localisation and lateralisation of activation in the auditory comprehension network during childhood and adolescence.

Study	*N*	Age range (years)	Language task	Baseline task	Process of interest	Core network (age independent regions)	Age-related increases	Age-related decreases
Ahmad et al., 2003	15	5-7	Passive listening to stories	Passive listening to task stimuli	**Semantic, syntactic, phonological and prosodic processing from**	LATERALISATION: Left lateralised activation in IFG, MFG, STG and IPL (LI values > 0.26).		
						LOCALISATION: Left STG, STS, MTG and AG		
Schmithorst et al., 2006	313	4-18	Passive listening to stories, including syntactically complex sentences	Passive listening to pure tones at unequal time intervals	**Semantic, syntactic, phonological, prosodic and complex auditory processing**	LATERALISATION: Left lateralised activation in ITG, IFG and MFG. Activation in other temporal regions, hippocampus, SFG and the AG was bilateral.		
						LOCALISATION: Using ICA identified six components including; bilateral STG, SFG, AG and hippocampus, and left medial temporal gyrus, IFG/MFG, IPL and precuneus/PCC.	LOCALISATION: Bilateral STG, left>right medial temporal gyrus, IFG, MFG and IPL.	LOCALISATION: Bilateral AG and left precuneus/post-central gyrus.
Wilke et al., 2005	15	6-14	*Beep Stories* task: Vigilent listening to auditory stories, in which 5-6 nouns were replaced by a sinus tone	Passive listening to sinus tones of varying lengths, played at different intervals	**Semantic, syntactic, phonological, prosodic and complex auditory processing**	LATERALISATION: left lateralised activation in a frontal lobe ROI.		
						LOCALISATION: Left IFG, DLPFC and SMA/pre-SMA, anterior STG/STS, posterior temporo-parietal areas.		
Plante et al., 2006	205	5-18	Vigilent listening to unintelligible sentences (preserve syntactic prosody). Participants were instructed to press the button when they heard a target sentence, based on it's syntactic prosody	Vigilant listening to pure tones. Participants were instructed to press the button when they heard a target tone	**Prosodic processing, complex auditory processing**	LATERALISATION: Bilateral activation for frontal ROI (LI = -0.08) and temporal ROI (LI=0.12), respectively.		
						LOCALISATION: Whole-brain activation not reported. ROI analysis showed activation in a bilateral frontal ROI (IFG, MFG, precentral gyrus) and temporal ROI (STG, MTG, ITG). Activation was greater in the temporal ROI then the frontal ROI.	LOCALISATION: Bilateral STG.	
Balsamo et al., 2006	23	5-10	Vigilent listening to auditory sentence (semantic category, e.g "things you can eat") and list of nouns. Participants intructed to press the button when a noun belongs to the category	Vigilent listening to reversed versions of noun task stimuli. Participants were instructed to press a button when they heard a pure tone.	**Semantic, syntactic, phonological, prosodic and complex auditory processing, and semantic decision making**	LATERALISATION: 78% of children had left lateralised activation across the whole brain.		
						LOCALISATION: Left IFG (BA 44+47), MTG, fusiform gyrus, medial frontal gyrus and SMA.		
Brauer & Friederici 2007	12	5-6	Vigilent listening to sentences which were syntactically or semantically incorrect. Participants were asked to press a button at the end of each sentence to indicate whether the sentence was correct or incorrect.	Vigilent listening to sentences which are semantically and syntactically correct. Participants were asked to press a button at the end of each sentence to indicate whether the sentence was correct or incorrect.	**1.Syntactically incorrect > correct = syntactic processing 2.Semantically incorrect > correct = semantic processing**	LATERALISATION: Both conditions = activation in the aSTG was bilateral. Syntactic condition = ctivation in the frontal operculum was bilateral, mSTG was left lateralised, SMG was right lateralised, and pSTG was only active in the left hemisphere. Semantic condition = activation in the mSTG was bilateral, in the pSTG was right lateralised, and lateral IFG, FO and SMG were only active in the left hemisphere.	LATERALISATION: Group comparison suggested a trend for increased lateralisation in the temporal lobe in adults versus children (assessed by comparing temporal lobe cluster sizes between hemispheres).	
	13	21-30				LOCALISATION: Both conditions = Bilateral STG (extensive), lateral IFG, FO and SMG. Syntactic condition = additional recruitment of left IFG. Semantic condition = no unique activation.		
Yeatman et al., 2010	14	10-16	Vigilent listening to syntactically complex auditory sentences followed by a picture stimuli. Participant asked to press a button if the picture accurately depicted the sentence.	Vigilent listening to syntactically simple auditory sentences followed by a picture stimuli. Participant asked to press a button if the picture accurately depicted the sentence.	**Syntactic processing from auditory input**	LATERALISATION: Bilateral activation in STG and MTG.		
						LOCALISATION: Left MTG, right STG and right insula.		LOCALISATION: Right fusiform gyrus and MFG.
Berl et al., 2010	59	4-12	Vigilent listening to stories. Participants were instructed to press a button when they heard a tone, presented at the end of each sentence. (4 developmentally appropriate difficulty levels).	Vigilent listening to reversed speech versions of task stimuli. Participants were instructed to press a button when they heard a tone, presented at the end of each sentenc.	**Semantic, syntactic, phonological and prosodic processing from auditory input**	LATERALISATION: Left lateralised activation was seen for 81% in Broca's area (IFG), 92% in Wernicke's area (STG+MTG+AG) and 73 % in the MFG.	LATERALISATION: Increasing left lateralisation of activation in Wernicke's area (STG+MTG) with age (trend level).	
						LOCALISATION: Bilateral STS (left>right), left IFG (BA 45+47), left orbitofrontal cortex (BA 11).	LOCALISATION: Frontal cortex (IFG and MFG), although only at trend level; active in 41% of 4-6 year olds,and >70% of 7-12 year olds.	
Lidzba et al., 2011	33	6-24	*Beep Stories* task: Vigilent listening to auditory stories, in which 5-6 nouns were replaced by a sinus tone	Passive listening to sinus tones of varying lengths, played at different intervals	**Semantic, syntactic, phonological, prosodic and complex auditory processing**	LATERALISATION: Bilateral activation in a fronto-temporal ROI, right lateralised activation in the cerebellum.		
						LOCALISATION: Bilateral STG and MTG, left AG, IFG and precentral gyrus, SMA and right cerebellum.	LOCALISATION: Right STG, bilateral MTG.	
Berl et al., 2012	57	4-12	Vigilent listening to auditory sentence descriptions of objects and animals (e.g. "A long yellow fruit is a banana"). Participant asked to press a button when the description was true. (4 developmentally appropriate difficulty levels)	Vigilent listening to reversed versions of the task stimuli, with a sinus tone inserted at the end of several stimuli. Participants asked to press a button when they heard the tone.	**Semantic, syntactic, phonological and prosodic processing, and semantic decision making**	LATERALISATION: Left lateralised activation was seen for 81% of participants in IFG, 88% in Wernicke's area (STG+MTG+AG) and 68% in the MFG. Activation in the cerebellum was right lateralised for 75% of participants.	LATERALISATION: Increasing left lateralisation of activation in Broca's area (IFG) and Wernicke's area (STG+MTG). Age accounted for 10% and 8% of change in Broca and Wernicke LI values, respectively (including when controlling for task performance).	LATERALISATION: Age-group comparisons showed activation in the cerebellum became more right lateralised with age.
						LOCALISATION: Bilateral (left>right) inferior frontal gyrus (bilateral BA 44 and 45, right BA 47), bilateral MFG and STG/STS, left SFG and fusiform gyrus, right cerebellum, bilateral (left>right) caudate, putamen and thalamus.		LOCALISATION: PCC, ACC and right IFG (age group comparisons showed this was driven by the youngest group: 4-6 years).

Passive tasks
Active tasks

Positive findings (effect of age)
Confounded by task performance
Negative findings (no age effect)
Analysis not performed/reported

Table 27: (overleaf) A review of fMRI studies investigating sentence-level auditory comprehension in children and adolescents. For each study lateralisation and localisation of activation in the core network (regardless of age) is described, followed by descriptions of age-related increases and decreases in network activation and lateralisation. Abbreviations include: superior, middle and inferior frontal gyrus (SFG, MFG, IFG), superior, middle and inferior temporal gyrus (STG, MTG, ITG), inferior parietal lobe (IPL), angular gyrus (AG), posterior and anterior cingulate cortex (PCC and ACC), frontal operculum (FO), dorsolateral prefrontal cortex (DLPFC), supplementary motor area (SMA), Brodmann area (BA), region of interest (ROI), laterality indices (LI), and also anterior, middle and posterior (a-, m-, p-). Colour coding has been used to differentiate between passive and active auditory comprehension tasks and between different types of findings (key below table).

3.1 Localisation of activation in the auditory comprehension network

3.1.1 Age-independent network

Three types of task have been used to investigate auditory comprehension in paediatric populations. Early studies compared passive listening to stories to either reversed speech (Ahmad et al., 2003) or pure tones (Schmithorst et al., 2006). More recent studies have investigated auditory comprehension during vigilant listening to stories; either with pure tones at the end of each sentence which need to be acknowledged with a button-press response - inducing active task monitoring (Berl et al., 2010) - or with pure tones replacing key nouns in the story - inducing word retrieval (Lidzba et al., 2011; Wilke et al., 2005). These tasks aim to encourage continued engagement with the task throughout scanning. Finally, in an attempt to measure task performance accuracy several studies have used decision-based language tasks, which require button-press (yes/no) responses, such as: deciding if nouns belong to a semantic category (Balsamo et al., 2006), judging if a sentence is semantically or syntactically correct (Brauer & Friederici, 2007), or if a heard sentence and picture are congruent (Yeatman et al., 2010). Compared to passive listening to pure tones, auditory comprehension induces bilateral (left>right) activation in the temporal lobes (STG, STS, MTG

and fusiform gyrus), as well as left AG, left IFG/MFG and precuneus/posterior cingulate cortex, bilateral SFG and hippocampus, precentral gyri and SMA (Ahmad et al., 2003; Lidzba et al., 2011; Schmithorst et al., 2006; Wilke et al., 2005). Vigilant and decision-making comprehension tasks show additional activation in frontal regions not seen in passive listening tasks, such as DLPFC (Wilke et al., 2005), left orbitofrontal cortex (Berl et al., 2010) and medial frontal gyrus (BA 8) (Balsamo et al., 2006). These regions are most frequently associated with executive functioning (Blakemore & Choudhury, 2006). All of the studies reviewed here compared auditory comprehension to a relatively low-level baseline task, such as passive or vigilant listening to pure tones or reversed speech. These low-level contrasts control only for basic sensory-motor aspects of task performance (for example, even reversed speech fails to control for prosody). As such, these studies are not able to distinguish between semantic, syntactic, phonological, prosodic and complex auditory processing, nor working memory or cognitive control aspects of task performance. Only two studies performed higher-level contrasts, with the aim of localising syntactic processing in children and adolescence. In these studies activation in left MTG, right STG and right insula was associated with comprehension of complex versus simple syntax (Yeatman et al., 2010), while left IFG was associated with syntactic versus semantic error detection (Brauer & Friederici, 2007). Although core semantic processing regions have not been localised in children, findings from all of the studies reviewed here do suggest regions associated with semantic processing in the mature language network (Price, 2012) are also engaged during auditory comprehension in children and adolescents. These regions include: bilateral/left lateral and medial temporal regions (STG, STS, MTG, fusiform gyrus), the left frontal lobe (IFG, MFG, SFG), and the inferior parietal lobe (AG) (Ahmad et al., 2003; Berl et al., 2010; Lidzba et al., 2011; Schmithorst et al., 2006; Wilke et al., 2005). In addition to these regions, children and adolescents show activation in the hippocampus, posterior cingulate cortex/precuneus, right SFG, left medial frontal gyrus, DLPFC and orbitofrontal cortex. These additional regions may be related to non-linguistic aspects of task performance however, such as memory and executive functioning. My own work in Chapter 5 suggests semantic and syntactic aspects of auditory comprehension are specifically associated with

activation in left IFG and bilateral temporal regions (anterior STS/STG, occipitotemporal cortex, basal temporal language areas and the temporal pole), regardless of age.

3.1.2 Age-related changes

Eight of the 10 studies reviewed here investigated age-related changes in localisation of activation during auditory comprehension. Age-related increases in activation were identified in four studies (50%), including: bilateral STG/MTG (Lidzba et al., 2011; Plante, Holland, & Schmithorst, 2006; Schmithorst et al., 2006), IFG/MFG (Berl et al., 2010; Schmithorst et al., 2006), as well as the fusiform gyrus and IPL (Schmithorst et al., 2006). Age-related decreases in activation were identified in only three studies (38%), including right frontal regions (IFG/MFG) and the fusiform gyrus (Berl et al., 2014; Yeatman et al., 2010), midline regions (anterior and posterior cingulate, and left precuneus) (Berl et al., 2014; Schmithorst et al., 2006), and bilateral AG (Schmithorst et al., 2006). Because these studies lack precision in localising different linguistic processes, it is not possible to infer the functional correlate of these changes. However, a few studies have begun to use higher-level contrasts to localise age-related changes in specific networks supporting auditory comprehension. Prosodic processing, for example, shows increasing activation in bilateral STG with age (Plante, Holland, et al., 2006). Another study suggests specialisation of left BA 44 for syntactic processing emerges late in childhood (>11 years of age) (Nunez et al., 2011), although groups in this study had small sample sizes (N≤10). In this chapter I aim to define age-related changes in the auditory comprehension network specifically associated with semantic and syntactic processing. In doing so I hope to contribute to our understanding of the development of functional ontologies in the language system, similar to work previously conducted in adults.

3.2 Lateralisation of activation in the auditory comprehension network

3.2.1 Age-independent network

In the pre-surgical setting language lateralisation is usually assessed with fMRI using single-word generation tasks, such as verb generation. Single word

generation tasks are specifically selected for pre-surgical assessment because they typically evoke a strongly lateralised response in the brain, and show good correspondence with invasive measures of language dominance (Arora et al., 2009; Baciu, Watson, et al., 2005; Benson et al., 1999; FitzGerald et al., 1997; Giussani et al., 2010; Liegeois et al., 2002; Roux et al., 2003; Suarez et al., 2009). Less is known about the correspondence of receptive language network lateralisation and either invasive measures or post-surgical outcome. As a first step towards validating the Listen and Name game for pre-surgical assessment I aim to characterise laterality of activation during performance of this task in healthy children, and its change with age. The majority of studies reviewed here suggest activation in the developing auditory comprehension network – regardless of age - is bilateral, but left lateralised (Ahmad et al., 2003; Balsamo et al., 2006; Berl et al., 2010; Berl et al., 2014; Wilke et al., 2005). There is some discrepancy between studies however, potentially due to variations in the regions and methods used to define lateralisation, as well as differences in tasks. Indeed, studies provide evidence for an effect of region on lateralisation, such that activation is more left lateralised for frontal versus temporal regions (Berl et al., 2014; Lidzba et al., 2011; Schmithorst et al., 2006). Studies have also shown an effect of task, such that activation in the temporal lobes is more left lateralised for syntactic versus semantic processing (Brauer & Friederici, 2007). Based on these studies, I hypothesise activation in classic Broca's and Wernicke's areas to be only weakly left lateralised. Little is known about lateralisation of activation in the temporal poles and basal temporal language areas during development. This is particularly surprising as these regions are implicated in core semantic processing (Binder et al., 2009), appear critical to language during intracranial mapping (Davies et al., 1994; Krauss et al., 1996; Luders et al., 1991), and are implicated in surgery for drug-resistant epilepsy for children (Harvey et al., 2008). Based on adult models of the semantic system (Binder et al., 2009; Lambon Ralph, Pobric, & Jefferies, 2009) I hypothesise bilateral activation in the temporal poles, and left lateralised activation in basal temporal regions, such as the ventral occipito-temporal cortex (Price, 2012).

3.2.1 Age-related changes

Of the 10 studies reviewed here, five investigated age-related changes in lateralisation; three (60%) provide evidence for age-related increases in lateralisation to the left hemisphere – particularly in temporal regions (Berl et al., 2010; Berl et al., 2014; Brauer & Friederici, 2007). However, as discussed in section 2.1, this may reflect changes in non-linguistic (e.g. sensory-motor or executive), as well as linguistic, processing. Lateralisation of activation for core linguistic processes, localised by high-level contrasts, is not well understood. In order to improve the role of fMRI in the paediatric neurosurgical decision making process, it is important to know if lateralisation of core language processes is more or less predictive of post-surgical language outcome compared to measures which take into account both linguistic and non-linguistic processing. I aim to provide an initial exploratory investigation of the effect of contrast (high-level versus low-level) on measures of lateralisation in this chapter, as a first step towards answering this important clinical question.

3.3 Summary: Localisation and lateralisation of activation

The results of the literature review summarised above provide evidence for a widespread auditory comprehension network engaged by children and adolescents - regardless of age, and consistent between studies. This network closely resembles the auditory comprehension network in the mature language system (Price, 2012) and is generally bilateral (with activation weakly lateralised to the left hemisphere). Age-related changes in this network are less well defined however. There is currently no clear evidence to suggest either increases or decreases in network activation are the norm. Instead, my review has highlighted the inconsistency of findings between studies, and the need for further work in this area. Crucially, many of the studies reviewed here share a basic confound which has not yet been discussed; task performance.

3.4 Task performance confounds on age-related changes in network activation

Many of the studies discussed above used a one-size-fits-all approach to task design. As such, younger participants may find the task more difficult than older participants – especially when children are compared to adults. This can result in age-related

differences in task performance accuracy. Up to half of apparent age-related changes in language network activation may be due to task performance differences (Brown et al., 2005). For example, improved performance on a semantic categorisation task is associated with increased activation in the left middle frontal gyrus and fusiform gyrus (Balsamo et al., 2006). Although no evidence could be found for an effect of in-scanner performance on lateralisation of activation during auditory comprehension (Berl et al., 2014), this study used age-appropriate tasks which aimed to control for such differences. None of the other studies reviewed here investigated the effects of in-scanner task performance on network activation. Of even more concern, some studies which reported age-related changes in network activation did not report control over task performance, even when performance accuracy differed by age (Plante, Holland, et al., 2006; Schmithorst et al., 2006). In one study, children actually performed 40% fewer trials relative to adults as part of the experimental design (Brauer & Friederici, 2007). These findings suggest some of the age-related changes in network activation reported in sections 2.1.1 and 2.2.1 may actually reflect age-related differences in task performance accuracy. Even in cases where performance accuracy is equivalent across age, task performance may have been considerably easier for older participants versus younger participants; introducing age-related differences in effort. This is concerning as task difficulty has been shown to effect language network activation. In children, the left MTG, right STG and right insula are more active when syntactic processing becomes more challenging for example (Yeatman et al., 2010). Some of the studies reviewed here compare age ranges as wide a 4-18 years (Schmithorst et al., 2006) or compare children to adults (Brauer & Friederici, 2007). In these studies all participants perform the same task. Many of the age-related changes reported in sections 2.1.1 and 2.2.1 may therefore be confounded by age-related differences in task difficulty, as well as performance accuracy. Compared to task performance accuracy (which can be assessed by the error rate), task performance difficulty is difficult to operationalise. Reaction time has been used as a proxy for difficulty in the naming literature (see Chapter 2, section 8.1). However, it is important to control for changes in simple reaction time (associated with processing speed and motor control rather than cognitive effort). Alternatively, fMRI tasks can be designed for different ability levels (Berl et al., 2010; Berl et al., 2014). Age-adjusted tasks aim to provide

control over both performance accuracy and difficulty. Interestingly, in contrast to studies which did not control for performance accuracy and task difficulty, studies which used age-adjusted tasks found no evidence for age-related increases in activation in the auditory comprehension network. These studies found only decreases in activation, in midline regions and the non-dominant frontal lobe (anterior and posterior cingulate cortices and right inferior frontal gyrus) (Berl et al., 2014). These regions were only recruited by very young participants (aged 4-6 years). Crucially, the authors showed age actually accounts for only a small amount of variation in language lateralisation in frontal and temporal regions (~10%). By using age-appropriate tasks and analysing only correct response-trials, I aim to identify age-related changes in the auditory comprehension network in children and adolescents (aged 5-16 years) whilst providing stringent control over task difficulty and performance accuracy.

3.5 Hypotheses

The studies reviewed here show that while brain regions sensitive to auditory comprehension have been relatively well defined in children and adolescents, little is known about the development of specific structure-function relationships. There is little consistent evidence, therefore, on which to base my hypotheses for expected changes in the semantic and syntactic processing network with age. However, several frameworks have been proposed which aim to account for functional-neuroanatomical changes associated with cognitive development (Johnson, 2001). These frameworks provide specific hypotheses for the development of the language system, and are summarised in Table 28.

Framework	General predictions	Specific predictions for age-related changes in fMRI activation in the auditory comprehension network	Summary of expected findings
Maturation	Changes in activation will be seen in brain regions which mature between 5-16 years of age.	• Age-related increases in activation in the temporal lobe (specifically in the STG/STS), inferior parietal lobe and dorsolateral prefrontal cortex, which mature up to 16 years of age (Lenroot & Giedd, 2006). • Activation will become more widespread across the perisylvian network with age, due to the protracted development of frontal, temporal and parietal association fibres (Lebel et al., 2008).	**Increasing activation** throughout the perisylvian network with age, and in prefrontal regions.
Skill Learning	Changes in activation will be seen for language skills which become increasingly proficient between 5-16 years of age.	• Because the ability to comprehend grammatically-simple sentences has been acquired by 5 years of age in typically developing children (Bates, 1997), age-related changes in activation will be minimal in the network supporting comprehension of simple sentences in children aged 5-16 years.	**Minimal/no changes** in network activation with age.
Interactive specialisation	Network activation will become refined with age, from widespread to a network of localised 'core' regions.	• Increasing activation in regions associated specifically with sentence-level semantic and syntactic processing from phonological input in adults (Price 2012): bilateral superior temporal sulci, middle and inferior temporal gyri, left ventral occipito-temporal cortex, Pars Triangularis and Orbitalis, superior frontal gyrus, and bilateral angular gyrus. • Regions outside this network (seen when comparing LN>rest) will show decreasing activation with age.	**Increasing activation** in semantic/syntactic processing regions, and **decreasing activation** in other regions (e.g. prefrontal cortex).

Table 28: Hypotheses for changes in fMRI activation in the auditory comprehension network from 5-16 years, according to three frameworks for cognitive development described by Johnson (2001; 2011); maturation, skill learning, and the interactive specialisation. Each framework provides general predictions for brain development and, based on current research, specific predictions for the auditory comprehension network.

4. METHODS

4.1 Participants

Forty three healthy children were recruited and assessed as described in Chapter 3. Participants were excluded according to the following criteria 1) in-scanner movement >1 voxel during either the task or baseline condition (N=7). Descriptives for the remaining sample (N=36, 5-16 years) and results of age-group comparisons are presented in Chapter 5 (section 4.3.3 and Table 18). To summarise, age groups were well matched for demographics, neuropsychological performance, practice time and mean in-scanner movement.

4.2 Neuropsychological assessment

Neuropsychological performance was assessed as described in Chapter 3. I compared age groups based on several verbal and nonverbal measures, all of which were well-matched across age (these variables and comparisons are summarised in Chapter 5, Table 18). Receptive language ability was assessed according to standard scores on the Recalling Sentences subtest of the CELF-IV or (for two participants with missing data) the *Concepts and Following Directions* (for 5-12 year olds) or *Semantic Relationships* (for 13-16 year olds) subtest from the CELF-IV. Performance on the *Recalling sentences* subtest is strongly correlated with performance on both the *Semantic Relationships* subtest (r=0.85, p=0.004) and *Concepts and Following Directions* subtest (r=0.47, p=0.03). These subtests represent comparable age-appropriate measures of sentence-level auditory comprehension similar to the Listen and Name game, and requires maintenance of verbal information in working memory whilst an action or decision is made.

4.3 Functional MRI

Functional MR images were acquired and pre-processed as per Chapter 3, with the exception that in-scanner movement was corrected using the motion fingerprint method (Wilke, 2012), as described in Chapter 4. All participants performed the Listen and Name Game (LN) and Alien Game (AG) as part of a larger neuroimaging protocol, as described in Chapter 3. The order in which LN and AG were performed was counterbalanced across participants.

4.4 In-scanner task performance

Task performance data are presented in Chapter 5 (section 4.3.3 and Table 18). To summarise, performance on the LN task was not related to age. AG performance accuracy improved with age. However the omission rate did not differ across age, suggesting all participants made an equivalent number of responses (whether correct or incorrect). Because just attempting to differentiate the gender of the speaker involves auditory processing and word retrieval, age-related differences in accuracy should not influence activity in regions of interest for this task. Listen and Name and AG did not differ in terms of practice time or in-scanner movement.

4.5 Localisation of the auditory comprehension network: Group map analysis

Individual participants' *t* contrast images were entered into one-samples t-test at the second-level, to identify regions of activation associated with auditory comprehension (LN>rest), sensory and motor components of task performance (AG>rest) and task-specific regions associated with sentence-level semantic and syntactic processing (LN>AG), regardless of age. Results of this group map analysis are described in Chapter 5, section 4.5.3.

4.6 Age related changes in network activation: Linear regression

At the second-level, two linear regression models were calculated with age as the covariate of interest to identify; 1) regions of the auditory comprehension network (LN>rest) associated with age, and 2) regions of the semantic and syntactic processing network (LN>AG) associated with age. For the auditory comprehension network (LN>rest) results were investigated and corrected throughout the whole brain, as age-related changes were hypothesised in regions beyond classic language cortex (associated with cognitive control aspects of task performance). For the semantic and syntactic processing network (LN>AG) age-related changes were hypothesised within regions engaged during auditory comprehension. Results of the linear regression for contrast LN>AG were therefore viewed (and corrected for multiple comparisons) within a mask of the auditory comprehension network (based on significant clusters of activation for contrast LN>rest, when $p<0{,}001$ and $k>10$ uncorrected).

4.6.1 Eigenvariate analysis

The first eigenvariate was extracted from age-dependent regions (clusters showing a significant relationship to age in the above linear regression model). The first eigenvariate represents the typical response of a cluster to the task, for each participant. Eigenvariate data for each region were entered into SPSS to inspect scatterplots for outliers. I also performed multiple linear regression analyses in SPSS to investigate whether age predicted activation in each region once providing statistical control over task performance accuracy and receptive language ability.

4.7 Lateralisation of the auditory comprehension network: Laterality indices

Laterality indices (LIs) were calculated using the LI toolbox, which uses a bootstrap approach to calculate a weighted mean from across statistical thresholds (Wilke & Schmithorst, 2006), as described in Chapter 1. Laterality indices were calculated in classic language ROIs: Broca's area (inferior frontal gyrus) and Wernicke's area (middle and superior temporal gyri), comparable to previous studies (Berl et al., 2014). Based on group map activation (described in Chapter 5) laterality indices were calculated in two additional ventral language ROIs: a basal temporal ROI (including the inferior temporal and fusiform gyri) and a temporal pole ROI (including the middle and superior temporal poles). Little is known about laterality of activation in these regions during language processing, despite their vulnerability to surgical resection in paediatric epilepsy (Harvey et al., 2008). For similar reasons LIs were also calculated for a medial temporal ROI (including the hippocampus and parahippocampal gyrus). Finally, laterality indices were calculated in a frontal ROI (including middle frontal gyri/dorsolateral prefrontal cortices, orbitofrontal and anterior cingulate cortex) in order to test the hypothesis that age-related changes in laterality occur predominantly in executive control regions. Laterality indices in the medial temporal and frontal ROIs were only calculated for the contrast LN>rest, as these processes are controlled for in the contrast LN>AG. All ROIs are displayed in Figure 59.

4.8 Age related changes in lateralisation of activation

4.8.1 Language regions

A mixed ANOVA was performed to investigate the effect of region (within-subject factor), baseline (within-subject factor) and age (between-subject factor) on laterality indices to test the hypotheses that age-related changes in laterality differed for high-level and low-level contrasts (LN>AG versus LN>rest), and between regions in the language network. This analysis was performed on LIs from language regions active in both contrasts (Broca, Wernicke, temporal pole and basal temporal ROIs).

4.8.2 Non-linguistic regions

One-way ANOVA and correlation analyses were performed to investigate the effect of age on LIs from the medial temporal and frontal ROIs.

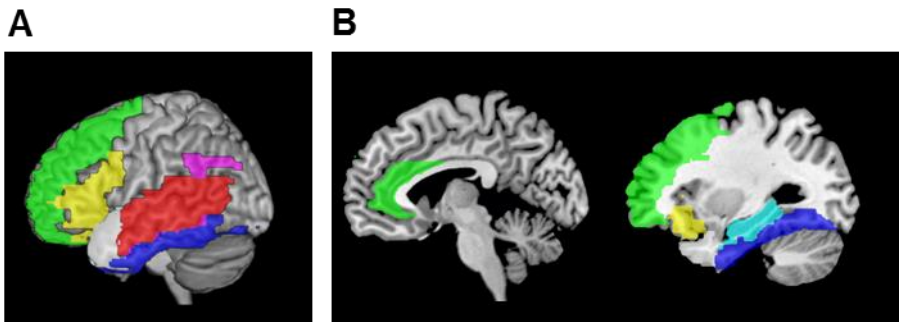


Figure 59: Regions of interest used for calculation of laterality indices shown on the rendered surface of the brain (A) and on sagittal slices in the left hemisphere (B). Regions include classic Broca's area (yellow) and Wernicke's area (red), as well as the temporal pole (white) and basal temporal cortex (blue). Laterality indices were also calculated in the medial temporal lobe (cyan) and in a frontal region of interest (green).

5. RESULTS

5.1 The auditory comprehension network

Results of the group map analysis are explained in detail in Chapter 5, section 4.5.3.

5.2 Age-related changes in localisation of activation

MNI co-ordinates, *z* and *p* values for significant local maxima (>8mm apart) are listed in Appendix 12.

5.2.1 Listen and Name (>rest)

Age-related increases in activation were seen in left Pars Orbitalis (BA 47), right fusiform gyrus and parahippocampal gyrus ($p<0.05$, FWE corrected). At a more lenient threshold ($p<0.001$, k>10 uncorrected) age-related increases in activation were seen in bilateral parahippocampal/fusiform gyri. A large cluster showed age-related increases in activation in bilateral frontal lobes, including: Pars Orbitalis (BA 47), precentral gyrus, medial superior frontal gyrus, SMA and around the inferior frontal sulcus. Additional increases in activation were seen in the right mid-cingulate cortex. Age-related increases were seen in bilateral temporal lobes (mid-STS/STG) and specifically in the right temporal lobe (anterior inferior temporal sulcus and primary auditory cortex). Activation increases were also seen in the left anterior insula and lingual gyrus. There were no age-related decreases in activation.

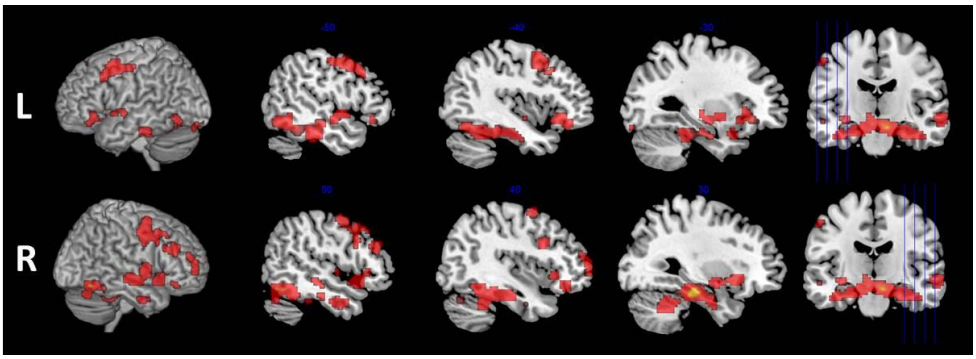


Figure 60: Age-related increases in activation in regions engaged during auditory comprehension of sentences. Results from linear regression analysis of age-related changes in activation for the comparison Listen and Name game (LN) > rest. Results are shown at two thresholds; $p<0.05$ FWE corrected (yellow) and $p<0.001$, k>10 uncorrected (red).

5.2.2 Listen and Name > Alien Game

Within the semantic processing network, the right ($p<0.05$, small volume corrected) and left ($p<0.001$, k>10, uncorrected) vOT showed increasing activation with age, as well as the left posterior inferior temporal gyrus ($p<0.001$, k>10, uncorrected). To ensure this did not reflect increasing inhibition of these regions with age in the Alien Game condition, I also performed linear regression to investigate age effects for the Alien Game. This analysis showed no effect of age on activation in vOT or left posterior inferior temporal gyrus, even at a lenient threshold ($p<0.01$, $k>10$).

5.2.2.1 Eigenvariate analysis

The first eigenvariate was extracted from each of the three age-dependent clusters identified in the linear regression analysis (Figure 61): right mid-fusiform (x=39, y=-37, z=-19; k=42), left mid-fusiform (x=-36, y=-37, z=-19; k=35), and left posterior inferior temporal gyrus (x=-48, y=-67, z=-10; k=12). Multiple linear regression analyses performed in SPSS showed age accounted for 43% of variation in the right mid-fusiform gyrus ($R^2_{adjusted}$=.43, $F(1,32)=25.50$, $p<0.001$) and 35% in the left mid-fusiform gyrus ($R^2_{adjusted}$=.35, $F(1,32)=19.08$, $p<0.001$) when controlling for task performance accuracy and receptive language ability. For the left posterior inferior temporal gyrus (pITG) multiple linear regression produced two models. The first model showed age accounted for 36% of variation in activation when controlling for task performance and receptive language ability ($R^2_{adjusted}$=.36, $F(1,32)=18.04$, $p<0.001$). The second model accounted for more variability (41%) and included both age and LN task performance accuracy as predictors of activation ($R^2_{adjusted}$=.41, $F(1,32)=12.63$, $p<0.001$). Partial regression plots showed activation in the left pITG increased with age, regardless of task performance accuracy. In contrast, activation decreased with more accurate performance of the LN task, regardless of age.

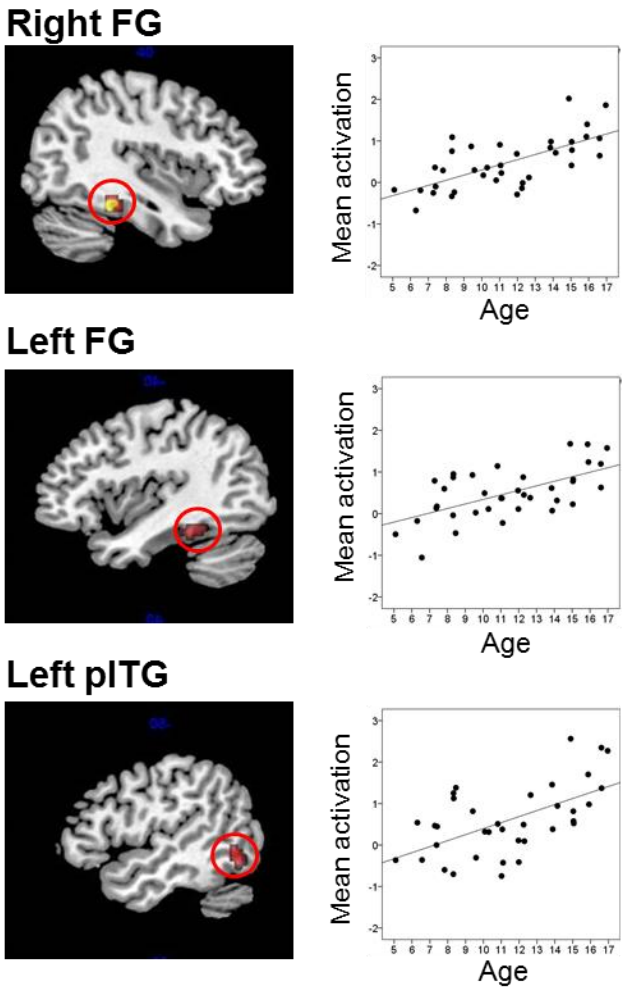


Figure 61: Age-related increases in activation within the semantic and syntactic processing network engaged during auditory comprehension. Results from linear regression analysis for the contrast LN>AG are presented on sagittal slices of the brain (left), showing age-related increases in activation for the right mid-fusiform gyrus (FG), left mid-fusiform gyrus and left posterior inferior temporal gyrus (pITG). The first eigenvariate is plotted against age for each region (right).

5.3 Age-related changes in network lateralisation

Laterality indices are displayed by region, contrast and age in Figure 62, and are listed in Table 29.

	TOTAL (N=36)				EARLY CHILDHOOD (N=11)				LATE CHILDHOOD (N=12)				ADOLESCENCE (N=13)				
	LN>rest		LN>AG		LN>rest		LN>AG		LN>rest		LN>AG		LN>rest		LN>AG		
REGION	**Mean**	SD	**Mean**	SD	**Mean**	SD	**Mean**	SD	**Mean**	SD	**Mean**	SD	**Mean**	SD	**Mean**	SD	**Classification**
Classic language regions																	
Broca	**0.26**	0.39	**0.26**	0.40	**0.27**	0.42	**0.22**	0.37	**0.31**	0.35	**0.26**	0.37	**0.21**	0.42	**0.29**	0.46	Left lateralised (weak)
Wernicke	**0.08**	0.37	**0.27**	0.36	**0.10**	0.39	**0.32**	0.39	**0.08**	0.36	**0.19**	0.31	**0.06**	0.38	**0.31**	0.39	Bilateral (left>right)
Language regions vulnerable to surgery																	
Basal temporal	**0.15**	0.43	**0.20**	0.38	**0.17**	0.58	**0.22**	0.45	**0.09**	0.34	**0.16**	0.36	**0.19**	0.38	**0.22**	0.35	Bilateral (left>right)
Temporal pole	**-0.13**	0.40	**-0.03**	0.32	**-0.11**	0.50	**0.01**	0.35	**-0.17**	0.41	**-0.01**	0.33	**-0.11**	0.31	**-0.09**	0.31	Bilateral (right>left)
Memory and cognitive control regions																	
Medial temporal	**0.11**	0.35	**-**	-	**0.04**	0.44	**-**	-	**0.11**	0.29	**-**	-	**0.16**	0.34	**-**	-	Bilateral (left>right)
Frontal regions	**0.21**	0.43	**-**	-	**0.19**	0.40	**-**	-	**0.21**	0.42	**-**	-	**0.23**	0.48	**-**	-	Left lateralised (weak)

Table 29: Laterality indices in classic language regions, language regions vulnerable to surgery in children with drug-resistant epilepsy, and in memory and cognitive control regions, for the whole sample and each age group. Laterality indices are shown for the low-level contrast condition (LN>rest) and the high-level contrast condition (LN>AG). The classification of lateralisation for each region is provided in the final column, where indices ≥0.2 are considered left lateralised, indices ≤-0.2 are considered right lateralised, and indices between -0.2 and 0.2 are considered bilateral, similar to previous studies (Berl et al., 2014).

5.3.1 Language ROIs: Mixed ANOVA

5.3.1.1 Effect of region

There was a main effect of region on lateralisation of activation ($F(3,99)=8.90$, $p<0.001$, $\eta^2=0.14$). For the low-level contrast condition (LN>rest), post hoc paired samples *t*-tests with Bonferroni correction for multiple comparisons ($p<0.004$) showed significantly lower laterality indices for the temporal pole relative to Broca's area and basal temporal regions (p values <0.005), and – at trend level – relative to Wernicke's area ($p=0.01$) (Figure 52). For the high-level contrast condition (LN>AG) laterality indices were again lower for the temporal pole relative to Broca's area, basal temporal regions, and Wernicke's area (p values <0.004) (Figure 52).

5.3.1.2 Effect of contrast

There was a main effect of contrast on lateralisation of activation ($F(1,33)=7.47$, $p=0.01$, $\eta^2=0.02$), such that laterality indices were higher (more left lateralised) for the high-level contrast condition (Mean=0.17, SD=0.04, CI=0.08 - 0.23) compared to the low-level contrast condition (Mean=0.09, SD=0.04, CI=0.008 - 0.17). There was no interaction between contrast, region, or age ($p=0.15$).

5.3.1.3 Effect of age

There was no main effect of age ($p=0.93$), and no interaction of age with either contrast or region (all p values >0.79). There was no correlation of age with LI values for any region (all p values >0.35).

5.3.2 Non-linguistic ROIs: One-way ANOVA

One-way ANOVA showed no significant difference in lateralisation of activation in medial temporal or frontal ROIs between age groups (p values >.73). Further, there was no significant correlation of LIs from these regions with age (p values >0.86).

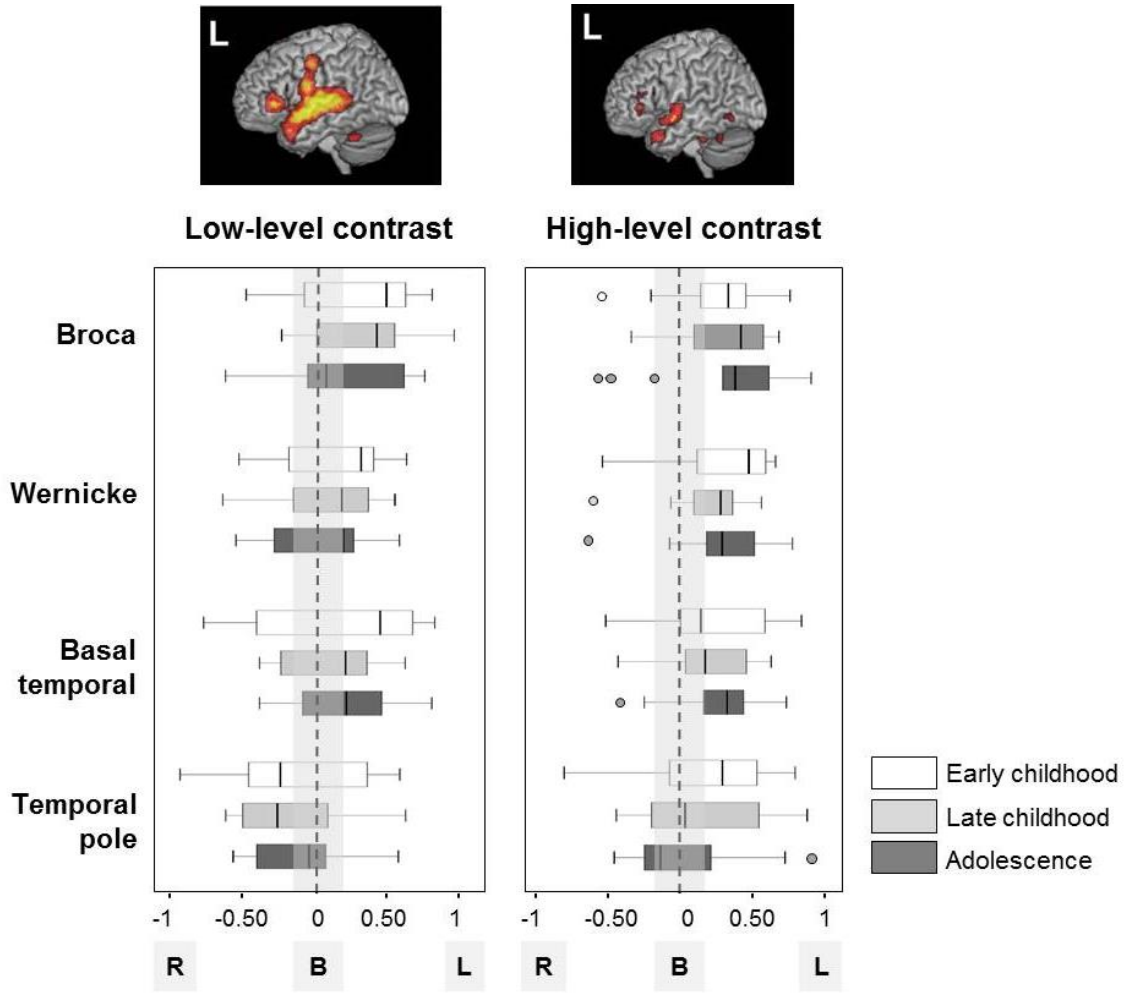


Figure 62: Laterality indices by region and age, shown for the whole auditory comprehension network (low-level contrast; left) and associated with semantic and syntactic processing specifically (high-level contrast; right). The low-level contrast condition shows laterality indices calculated using contrast LN>rest. The high-level contrast condition shows laterality indices calculated using contrast LN>AG. Outliers are depicted (dots). Data points within the grey band centred around 0 (dashed line) are categorised as bilateral (B). Data points to the right of the grey band are categorised as left lateralised (L), and data points to the left of the grey band are categorised as right lateralised (R).

6. DISCUSSION

6.1 The auditory comprehension network in children and adolescents

My findings suggest children and adolescents recruit a broad network of regions during auditory comprehension of sentences, similar to previous studies in children, adolescents and adults (Ahmad et al., 2003; Balsamo et al., 2006; Berl et al., 2010; Berl et al., 2014; Lidzba et al., 2011; Nunez et al., 2011; Plante, Holland, et al., 2006; Price, 2012; Schmithorst et al., 2006; Wilke et al., 2005; Yeatman et al., 2010). This network includes regions which, in adults, have consistently been associated with auditory processing (bilateral STG), semantic processing (bilateral STG and temporal pole), semantic retrieval (left Pars Triangularis and Orbitalis), auditory imagery (left planum temporale), breathing control during speech production (bilateral thalamus and left anterior insula), motor planning (SMA), orofacial motor activity (bilateral precentral gyri) (Price 2012), as well as other general motor processing regions (bilateral globus pallidus and left rolandic operculum). Activation was also seen in regions shown to be critical for naming using intracranial mapping (left vOT, or basal temporal language area; BTLA) (Davies et al., 1994; Krauss et al., 1996; Luders et al., 1989; Luders et al., 1986; Luders et al., 1991), as well as bilateral hippocampi. Finally activation was also found in the medial superior frontal gyrus, associated with cognitive control during task performance (Rushworth et al., 2004) and also proposed to form part of the core semantic system (Binder et al., 2009).

6.2 Semantic and syntactic processing regions in children and adolescents

In children and adolescents, comprehending a sentence involves regions which in adults have previously been associated with multimodal integration of semantic information (bilateral temporal poles), semantic association (anterior STS/STG), semantic retrieval (left Pars Triangularis and Orbitalis), semantic associations or visual expectations/imagery (bilateral ventral occipito-temporal cortex), naming and comprehension (left vOT), and articulatory recoding, phonological expectations and short term memory (left Pars Opercularis) (Price 2012). Activation was also seen in the left thalamus.

6.3 Age-related increases in activation in the auditory comprehension network

Previous investigations of age-related changes in the auditory comprehension network have often been confounded by age-related differences in task difficulty or accuracy (in at least ~30% of studies). Using age-adjusted tasks designed to control for task performance difficulty and accuracy, and by analysing only correct response trials, I investigated age-related changes in regions sensitive to auditory comprehension (versus rest). These regions may be involved in core language functions (e.g. semantic or syntactic processing), sensory-motor processing (e.g. hearing or articulation), or cognitive control aspects of task performance (e.g. sustained attention). Although these different functions cannot be distinguished using this low-level contrast, findings are interpreted in the context of previous structure-function mappings from the neuroimaging literature on language (Price 2012). Age-related increases in activation were seen in regions of the language network associated with semantic processing (bilateral BA 47 and mid-STG/STS) and sensory-motor processing (bilateral precentral gyri, supplementary motor area, left anterior insula and left lingual gyrus). Finally, increased activation was also seen bilaterally in relatively large clusters centred on the inferior frontal sulci. This area has been implicated in a cognitive control system (Sundermann & Pfleiderer, 2012) and – specifically - a syntactic working memory network (Makuuchi, Bahlmann, Anwander, & Friederici, 2009; Makuuchi & Friederici, 2013) which may be late maturing (Brauer et al., 2011; Friederici, 2012). These findings provide further evidence for age-related increases of activation in the inferior frontal gyrus and superior temporal gyrus, similar to previous studies (Berl et al., 2010; Plante, Holland, et al., 2006; Schmithorst et al., 2006). I also provide novel evidence for increasing activation in bilateral vOT (discussed below), sensory-motor regions, and frontal regions associated with working memory. These findings provide evidence to suggest increases in activation may be the norm for the typically developing auditory comprehension system; at least when grammatically simple sentences are used. Further, increases were not limited to core language areas, as predicted by the Interactive Specialisation framework. Instead, increases in activation were seen across the perisylvian network and in frontal regions, consistent with the maturation framework of cognitive development.

6.4 Age-related increases in semantic and syntactic processing regions

Previous investigations of age-related changes in the auditory comprehension network used relatively low-level contrasts. Consequently, we do not currently know how specific structure-function relationships develop. Without a thorough understanding of the specific conditions under which regions of the brain are and are not activated during development, the role of fMRI in the pre-surgical decision making process will remain constrained to assessments of language lateralisation, and not localisation. Using a high-level contrast, I investigated age-related changes in regions specific to semantic and syntactic processing from phonological input (i.e. regions increasingly active for the processing of meaningful sentences versus unintelligible sentences). Among the regions involved in semantic and syntactic processing, age-related changes were surprisingly focal, showing increases in activation in bilateral ventral occipito-temporal cortex and left posterior inferior temporal gyrus only. Each of these findings are discussed below:

6.4.1 Bilateral ventral occipito-temporal cortex

Activation in the left and right ventral occipito-temporal cortex (vOT) increased with age. Even when statistically controlling for task performance and receptive language ability, age accounted for 43% and 35% of variation in activation of the right and left vOT region, respectively. The clusters identified here were centred around Talaraich co-ordinates -32, -39, -19, which are slightly medial and anterior to the visual word form area 'proper' (Talaraich co-ordinates: x=-43, y=-54, z=-12) (Cohen et al., 2002). Cohen and colleagues propose this more anterior region supports amodal abstract representations, rather than abstract representations for written words specifically. Indeed, my co-ordinates are more consistent with those associated with semantic processing during reading, rather than phonological decoding of orthography (Seghier, Lee, Schofield, Ellis, & Price, 2008). Some authors even hypothesise this portion of the mid-fusiform gyrus is as important to semantic processing as the anterior temporal lobes (Binder & Desai, 2011; Mion et al., 2010).

6.4.2 Left posterior inferior temporal gyrus

In adults, the left posterior inferior temporal gyrus is most consistently associated with word retrieval from semantics, and is considered to support semantic access in the functional neuroanatomical model of language (Price, 2012). Together, age and task performance accuracy accounted for 41% of variation in this region (with age accounting for 36% of the variance alone). Activation of this region increased with age, and with more errors (poorer performance). Inaccurate responses may induce greater activation in this region due to more extensive semantic search or competition in an attempt to find the correct answer.

None of the studies investigating auditory comprehension reviewed in the introduction of this chapter provide supporting evidence for age-related changes in these regions (Table 27). This may be due to the use of low-level contrasts, lack of control over task performance, and differences in task design. Encouragingly however, extremely similar findings have been reported by a study of word-picture matching in a large sample of 283 children aged 5-18 years (Schmithorst, Holland, & Plante, 2007). This study used independent component analysis to extract functional networks supporting semantic access from both pictures (visual non-linguistic input) and heard words (phonological input) simultaneously. The authors found activation in the network supporting task performance increased with age; they found no evidence for age-related decreases in activation. Specifically, age-related increases were seen in bilateral fusiform gyri and posterior inferior temporal cortex. My results provide novel evidence for on-going maturation of similar regions during semantic processing, but from *sentence-level* phonological input. Crucially, Schmithorst and colleagues found age-related increases in activation were most evident for the right fusiform gyrus. In my analysis age-related changes in network activation only survived correction for multiple comparisons in the right fusiform gyrus, suggesting this is a robust finding. Based on the role of the right fusiform gyrus in picture naming (Etard et al., 2000), the authors suggest this may reflect increasing recruitment of a more efficient and automatic object recognition system to support picture naming with age. As discussed in Chapter 5, evidence from patients

with stroke suggests vOT supports amodal lexical access in adults (DeLeon et al., 2007). Consistent with this evidence, Price and Devlin propose the vOT supports direct semantic access from learned (familiar) visual patterns – including words, not just objects (Price & Devlin, 2011). My findings provide novel evidence to suggest this region also supports direct semantic access from learned (familiar) auditory patterns. Although the Alien Game was not optimised to control for visual imagery during sentence comprehension, this is typically associated with activation in the intraparietal sulcus, as opposed to vOT (Just, Newman, Keller, McEleney, & Carpenter, 2004). My findings therefore provide evidence for very focal increases in activation, in regions associated with semantic processing; consistent with predictions of the Interactive Specialisation hypothesis (Table 28).

6.5 Lateralisation

As hypothesised, and similar to previous findings in children and adolescents (Ahmad et al., 2003; Balsamo et al., 2006; Berl et al., 2010; Berl et al., 2014; Wilke et al., 2005), activation during auditory comprehension was weakly left lateralised in classic Broca's area, as well as in more 'executive' frontal regions (including middle frontal gyri/dorsolateral prefrontal cortex, orbitofrontal cortex and anterior cingulate cortex). Activation was bilateral (left>right) in Wernicke's area. Activation was significantly less left lateralised in the temporal poles, which showed bilateral activation with a right>left distribution (LI values ~-0.13). I therefore provide novel evidence suggesting the anterior temporal lobes form part of a bilateral semantic system during childhood and adolescence, as seen in adults (Lambon Ralph et al., 2009; Patterson, Nestor, & Rogers, 2007). These regions have been specifically related to storing amodal semantic representations in the mature language system (Binder et al., 2009). Activation in the basal temporal ROI (the ITG and fusiform gyrus, including vOT) was bilateral (left>right). In adults, activation in vOT is consistently and strongly lateralised to the left hemisphere (Cai et al., 2008; Seghier & Price, 2011, 2013). In contrast, right vOT is mostly known for its role in face processing (Kanwisher, McDermott, & Chun, 1997; McCarthy, Puce, Gore, & Allison, 1997). My findings provide novel evidence to suggest children and

adolescents recruit both the left and right vOT during auditory comprehension. Bilateral activation may suggest a lack of specialisation for language and faces, as proposed by the Interactive Specialisation framework (Johnson, 2001, 2011). However, laterality of activation in the vOT region did not increase with age, as might be expected. Indeed, there was no effect of age on lateralisation of activation in any of the regions included in this study. Other studies have also either failed to find an association between age and lateralisation (Lidzba et al., 2011), or suggest age accounts for only ~10% of variability in lateralisation (Berl et al., 2014; Holland et al., 2007). Delineating factors which influence language lateralisation will be crucial to improving the reliability and validity of fMRI assessments of language dominance in the pre-surgical setting. Such analyses may even highlight biomarkers of the potential for re-organisation of function, or may one day provide targets for interventions which aim to influence language lateralisation (for example, pre-surgically). In this study I was able to identify a methodological influence on lateralisation; contrast (or baseline). This analysis suggested core language processes (identified with high-level contrasts) are more strongly lateralised than supporting functions (identified with low-level contrasts).

7. CONCLUSIONS

Auditory comprehension is supported by a widespread perisylvian network with minimal change across childhood. Activation in this network is bilateral in the temporal lobes (particularly the temporal poles), and is weakly left lateralised in frontal regions. At most, age contributes only minimally to individual differences in lateralisation of the auditory comprehension network. Increases in activation appear to be the norm for development of this system. In line with the maturation framework, widespread increases of activation are seen in late-maturing frontal and temporal regions associated with sensory-motor processing and executive functioning. In contrast, core linguistic processing is supported by a limited set of focal brain regions, a sub-set of which show increasing activation with age; more consistent with the Interactive Specialisation framework. These regions include bilateral fusiform gyri (corresponding to the basal temporal language area or vOT) and the left posterior ITG.

Activation of these regions may reflect increasing cross-modal integration of semantic information, or the use of a more direct - automatic - route to semantics (similar to the lexical route in models of reading, and the object recognition system in picture naming). I have demonstrated the value of providing stringent control over task performance when investigating development of cognitive networks, and of using high-level contrasts to map the development of specific structure-function relationships. Crucially, these findings are of relevance to the pre-surgical decision-making process: The network supporting (simple) auditory comprehension is in place by at least 5 years of age, with on-going changes localised to focal regions at the base of the temporal lobes, which are relatively well-protected from surgery in the majority of cases. It does however highlight a degree of risk associated with removal of these regions for patients with a more posterior surgical target in occipito-temporal regions (approximately 2-23% of cases; Harvey et al., 2008).

GENERAL DISCUSSION

1. ADDRESSING THE AIMS OF THIS THESIS

This thesis had two main aims. First, I aimed to develop a panel of fMRI tasks optimised for pre-surgical planning, and optimised for use with young children and children with cognitive impairment. This new task panel was designed to improve pre-surgical assessments of language localisation (as opposed to language lateralisation). Second, I aimed to improve our understanding of typical language development by localising age-related changes in regions supporting semantic and syntactic processing, while providing stringent control over task performance accuracy and effort. Models of typical development are important in the pre-surgical setting to inform predictions of the cognitive risk posed by surgery performed in the context of development, and to contribute to decisions regarding the optimal timing for surgery. How well this thesis addresses each of these aims is outlined below.

1.1 Develop a new fMRI task battery to improve localisation of function in the pre-surgical setting

1.1.1 Functional MRI optimisation for the pre-surgical setting

I successfully designed and developed an fMRI task panel for investigating language, optimised for pre-surgical mapping:

1. The task panel provides a flexible toolkit for mapping different parts of the language system (Chapter 5) which can be tailored to the individual patient's needs (Chapter 6), by including both receptive and expressive language tasks, at both the word and sentence level, and in both visual and auditory modalities (Chapter 2).
2. The new task battery enabled mapping of the ventral language system (Chapter 5) often implicated in surgery (Chapter 6), through the use of sentence-level tasks and high-level sensory-baseline conditions.
3. Tasks were designed such that participants were actively engaged, while keeping executive function and working memory demands low. This

reduced identification of non-linguistic regions supporting task performance (Chapter 5).

4. Participants produced overt speech responses for each stimulus. This provided in-scanner performance monitoring, which could reduce up to 60% of scan failures in the pre-surgical setting (Chapter 1), and also increased the sensitivity of fMRI to detect language regions on an individual basis (Chapter 1).
5. The task panel enabled mapping of specific parts of the language network: 1) Brain regions supporting input and output to the language system could be localised using sensory-motor baseline tasks (Chapter 5, and Patient A in Chapter 6), 2) Regions supporting semantic and syntactic processing could be localised by comparing language tasks to their respective (high-level) sensory-motor baseline task (Chapter 5 and 6).
6. Candidate core language regions could be localised by comparing findings across tasks and modalities (Chapter 5 and Chapter 6, Patient F).
7. Tasks are compatible for use with intracranial mapping, allowing for validation of findings with invasive measures (Chapter 6, Patient A).
8. Tasks were developmentally appropriate (Chapter 2) and could be performed by children with cognitive impairment (Chapter 6).

1.1.2 Functional MRI optimisation for young children and children with cognitive impairment

The new task panel was also optimised for young children and children with cognitive impairment:

1. Scan failure rates caused by poor task performance were minimised (Chapter 5), due to:
2. Pre-scan preparation using a child-friendly protocol (Chapter 3)
3. Developmentally-appropriate tasks with four difficulty levels (Chapter 2). Participants can be matched to one of four difficulty levels depending on their age (Chapter 5) or their age-equivalent level of functioning (Chapter 6).

4. Participants were actively involved in task performance, and produced overt speech responses which provided a measure of in-scanner performance.
5. Tasks involved naturalistic use of language, and minimised executive load.
6. Developmentally appropriate design of tasks meant task difficulty was adjusted across age (Chapter 2), and task performance accuracy controlled across age (Chapter 5).

1.2 Contribute to models of typical language development

In the introduction to this thesis I highlighted the need for a more comprehensive understanding of how the language network develops, and how it may be affected by childhood onset epilepsy. This information is invaluable for making predictions in surgical candidates, regarding expected changes in language lateralisation and localisation with age, and in the context of on-going seizure activity. In the future, this information may be used to 1) inform decisions regarding the optimal time for surgery, 2) identify optimal candidates for surgery, and 3) predict language outcome. Two aspects of fMRI task design have so far limited our understanding of typical language development. I aimed to address both of these methodological limitations and, in doing so, contribute to models of typical language development. How I addressed each of these limitations is outlined below, along with my main findings:

1.2.1 Localising semantic processing within the language network

Previous studies of language in children have used fMRI tasks with relatively low-level baseline conditions, which are not adequately matched for sensory-motor or executive demands. Because these tasks identify activation associated with many processes, it has not been possible to investigate specific structure-function mappings in children. Models of typical development are therefore limited. Although several frameworks have been proposed which aim to account for and predict developmental changes in cognitive systems (Chapter 8), these cannot be formally tested with current methods. By using high-level sensory baseline conditions, I was able to distinguish between regions supporting task

performance, and regions supporting semantic processing from different input modalities (pictorial, orthographic and phonological). By identifying activation common across modalities, I was able to provide novel evidence in children for a core amodal semantic processing network, including bilateral anterior fusiform/parahippocampal gyri, left temporal pole, vOT, Pars Triangularis (BA 45) and left medial temporal cortex. These regions are similar to the semantic processing network reported in adults (Binder et al., 2009), and emphasises the integration of memory with linguistic processing.

1.2.2 Localising the effect of age on the language network

Previous studies of language development often used the same fMRI task over a broad age range, without accounting for age-related differences in performance difficulty or accuracy. The effects of age on the language system are therefore not well understood. The extent to which fMRI can inform decisions regarding the optimal timing for surgery (in terms of minimising risk to language function) will remain limited until the effects of age are better understood. I used the new task panel, designed and validated in the early parts of this thesis, to investigate typical language development (Chapter 8), while providing stringent control over task difficulty (Chapter 2) and performance accuracy (Chapter 5). I focused on the development of auditory comprehension as this engages the ventral language system, which is particularly relevant to language outcome in children with epilepsy (Chapter 7). This analysis produced three novel findings:

1.2.2.1 Prolonged development of semantic association areas in ventral temporal cortex.

Age-related changes in the semantic processing network were limited, and showed increasing activation in ventral temporal regions. Specifically, activation in the left posterior ITG and bilateral vOT (the mid-fusiform gyrus) increased with age. Functional MRI and PET studies in adults suggest these regions are involved in multimodal integration of semantic information (Binder, 2009), to gain direct semantic access to a word form (Price & Devlin,

2011). I provide novel evidence to suggest regions supporting direct semantic access, via the ventral language pathway, continue to develop into adolescence.

1.2.2.2 Classic language regions are established early.

My findings show Broca's and Wernicke's areas are engaged from as early as 5 years of age during comprehension, with little change across age when task difficulty and performance accuracy are controlled.

1.2.2.3 Children do not show consistent activation in the angular gyrus.

I found no evidence for involvement of the AG during picture naming, sentence generation, auditory comprehension or reading comprehension. Activation in the AG is often reported for the mature language system however, and has been associated with semantic decision-making (Demonet et al., 1992; Demonet, Price, Wise, & Frackowiak, 1994; Raettig & Kotz, 2008), language processing under increased semantic demands (Benson et al., 2001; Homae, Hashimoto, Nakajima, Miyashita, & Sakai, 2002; Noppeney & Price, 2002), top-down semantic predictions (Obleser, Zimmermann, Van Meter, & Rauschecker, 2007), and multimodal semantic integration (Binder et al., 2009). In children, activation is reported in the AG during verbal fluency (Karunanayaka et al., 2010; Szaflarski, Schmithorst, et al., 2006) and during comprehension when semantic retrieval demands are increased (Lidzba et al., 2011; Schmithorst et al., 2006; Wilke et al., 2005). The AG may not have been engaged by the tasks I developed in this thesis due to relatively low semantic demands (there was no semantic decision element to task performance, and stimuli were matched to each participant's ability level). Alternatively, this may reflect a genuine developmental effect. It is possible the AG matures relatively late, possibly associated with prolonged development of top-down predictions (Vigneau et al., 2006); this requires further investigation however.

2. SPECIFIC FINDINGS

My hypotheses for each Chapter in this thesis, whether they were supported, and specific evidence are listed in Table 30.

Hypotheses	Supported?	Specific results
Chapter 1: A systematic comparison of covert and overt speech in children using fMRI		
It is feasible to image overt speech when pre-processing methods are used to repair movement-related artefacts.	**Yes**	• A larger proportion of scans were rated as satisfactory/good quality when verb generation was performed overtly (76%) than covertly (55%).
Overt speech will induce stronger activation in core language regions.	**Yes**	• More participants showed activation in Wernicke's area during overt verb generation (92%) than during covert verb generation (68%). • There was a trend to suggest more participants showed activation in Broca's area during overt generation (90%) than during covert verb generation (66%).
Overt speech will induce activation in additional language regions versus covert speech, as a more valid model of expressive language.	**No**	• Relative to covert speech, overt speech was associated with increased activation in speech-motor regions only: primary and pre-motor cortex, ACC and midbrain.
Chapter 2: Design, development and pilot of a new fMRI task battery: 'The Panda Games'		
Children will perform less accurately compared to adults even on relatively simple language tasks, and will have slower reaction times.	**Yes**	• Children produced more errors than adults on all tasks. • Children had significantly longer RTs compared to adults on naming tasks (PN and RN/LN), and took longer to produce a sentence in the PD task.
Reaction times will be longer, and accuracy	**Yes**	• Between-level comparisons of pilot data (error rates and reaction

poorer, for more difficult levels of each task in the new battery, or children and adults.		time) confirmed both children and adults found levels 1-4 progressively more difficult for all three naming tasks in the Panda Games. • Between-level comparisons of pilot data (response accuracy and the *H* statistic) confirmed both children and adults found levels 1-4 progressively more difficult for PD.
Chapter 4: Identification and repair of speech related movement artefacts in children.		
Pre-processing methods which remove or de-weight time-points affected by large movements will be effective at removing speech-movement related artefacts.	**No**	• There was no significant difference in the prevalence of speech-movement related artefacts between four repair methods. • There was a trend to suggest the Robust Weighted Least Squares (rWLS) method reduced the prevalence of speech-movement related artefacts (LN=2%, PD=9%) compared to other methods (e.g. LN=14% and PD=28% for the standard GLM). This method de-weights noisy volumes during calculation of the beta-weights.
Methods which account for severe movement artefacts will be most effective at improving scan quality.	**No**	• Data quality was reduced by the robust weighted least squares method, due to reduced sensitivity to detect activation in language regions.
Chapter 5: Validation of the Panda Games in healthy children and adolescents aged 5-16 years.		
Scan failure rates will be lower compared to previous studies which did not use developmentally-appropriate tasks.	**Yes**	• Failure rates for the Panda Games were relatively low (ranging from 16% for PN to 42% for PD) compared to previous studies reporting failure rates between 31-57% (Byars et al., 2002; Yerys

		et al., 2009).
Movement will be the main cause of scan failure.	**Yes**	• The majority of scan failures were caused by in-scanner movement (and associated artefact), accounting for 100% (PN), 61% (PD), 82% (LN) and 50% (RN) of scan failures.
How participants felt during scanning (including anxiety and physical comfort) will be associated with in-scanner movement, and should be targeted by future research which aims to reduce scan failure rates in children.	**Yes (partially)**	• In-scanner movement was correlated with lower scores on a measure of *in-scanner experience* (rho=-.43, *p*=0.04) and *head-phone comfort* (rho=-.47, *p*=0.02). Lower scores indicate a less positive experience/more discomfort. • The in-scanner experience measure included items assessing physical comfort inside the scanner and enjoyment of the Panda Games. • Contrary to hypothesis, there was no effect of anxiety on in-scanner movement.
Mock scanning will reduce in-scanner movement, improve in-scanner task performance, and overall will reduce scan failure rates in children.	**Yes (partially)**	• In-scanner task performance was improved in children who underwent mock scanning (mean=93%) compared to participants who did not undergo mock scanning (mean=89%) despite equal practice time. • There was no significant difference in in-scanner movement or scan failure rates between participants who did and did not undergo mock scanning.
The new task battery will induce activation in	**Yes**	• The Panda Games identified activation in networks supporting

receptive and expressive networks, including regions predicted by models of the mature language system.		word generation, sentence generation, auditory comprehension and reading comprehension, similar to those hypothesised by the functional neuroanatomical model of language in the mature language system (Price, 2012).
Activation in language regions will be less left-lateralised in children, compared to predictions of adult models.	**No**	• Second-level t-tests showed activation in left hemisphere language regions, with little evidence for strong right hemisphere involvement.
Activation in frontal regions will be variable across tasks in children.	**No**	• Second-level t-tests showed activation in left Pars Triangularis (BA 45) and Pars Orbitalis (BA 47) in children as young as 5 years of age for every language task in the Panda Games battery (although only at a very low statistical threshold for PD). • Conjunction analysis showed consistent activation in left Pars Triangularis (BA 45) during naming regardless of input modality, in children as young as 8 years of age.
Activation will be seen across tasks in classic Broca's area (the inferior frontal gyrus), the inferior parietal lobe and the bilateral anterior temporal lobes, due to their involvement in amodal semantic processing.	**Yes (partially)**	• Conjunction analysis showed activation associated with amodal semantic processing in the left vOT, left temporal pole, left Pars Triangularis (BA 45), bilateral anterior fusiform gyrus and left entorhinal cortex. • Contrary to hypothesis, neither individual task group maps nor conjunction analysis showed activation in the inferior parietal lobe in children and adolescents.

Activation in Wernicke's area (the middle and superior temporal gyri) will be specific to comprehension tasks.	**Yes**	• Second-level t-tests comparing LN>rest showed activation in bilateral STG. • Second-level t-tests comparing RN>rest showed activation in left posterior MTG/ITG. • There was no significant activation in left STG or MTG during PN or PD.
Chapter 6: Pilot of the Panda Games in a representative sample of epilepsy surgery candidates aged 5-16 years		
Task performance on developmentally-appropriate versions of the new task panel will be improved compared to performance on a task which is not adjusted for ability level (verb generation).	**Yes (partially)**	• Performance accuracy was higher for PN (87%) than for verb generation (62%); both are single-word generation tasks. • Performance accuracy did not differ between verb generation and PD (68%) or LN (51%).
The new task panel will show increased sensitivity and specificity to detect activation in language regions compared to a task commonly used in the pre-surgical setting, which is optimised for assessing language lateralisation as opposed to localisation (verb generation).	**No**	• There was no significant difference between the Panda Games and verb generation in the number of participants who demonstrated significant activation in language regions (when $p<0.01$, $k>10$). • There was no significant difference between the Panda Games and verb generation in the sensitivity or specificity to detect activation in the language network (assessed by extent of activation in language regions versus a non-language control region).

The sensitivity of the new task panel to detect critical language regions (in relation to intracranial mapping) will be similar to findings from previous studies, which used lower-level baseline conditions.	**Yes**	• A case study (Patient A) showed the Panda Games have a similar level of sensitivity to detect critical language sites (0.75) to fMRI tasks with a lower-level baseline condition, reported in previous studies (~0.59-1.00).
Compared to values previously reported in the literature, the specificity of the new task panel to detect critical language regions (in relation to intracranial mapping) will be increased, due to the use of a high-level baseline condition.	**Yes**	• A case study (Patient A) showed the Panda Games had high specificity to detect critical language sites (0.70) relative to previous studies (~0.54).
The new task panel will provide flexibility in mapping of specific networks according to the individual patients' needs.	**Yes**	• Patient A and Patient F showed mapping of ventral and posterior temporal regions using PN and PD was beneficial in determining the risk posed by surgery to language function. • Patient F demonstrated the value of using a battery of language tasks, as opposed to a single task, to provide more confidence in findings (in this case showing consistent findings across two tasks; PN and PD). • Patient A demonstrated the value of having both visual and auditory tasks available (Patient A would not comply with auditory tasks) • Patient F demonstrated the value of using developmentally

		appropriate tasks (successful performance on PN and PD despite significant cognitive impairment).
The new task panel will allow identification of ventral language cortex which cannot be mapped using fMRI tasks currently used in the pre-surgical setting (which have been optimised for language lateralisation rather than language localisation).	**Yes (partially)**	• Patient F demonstrated that - on an individual level - the Panda Games can identify language cortex which may not be activated during performance of fMRI tasks currently used in the pre-surgical setting (verb generation). • Contrary to hypothesis – at the group level - there was no significant difference in the proportion of patients showing activation in ventral language regions between the Panda Games and an fMRI task currently used in the pre-surgical setting (verb generation).
Chapter 7: A systematic comparison of functional activation and functional connectivity methods for investigating language development in healthy children and children with epilepsy		
Functional activation will be equivalent between healthy children and children with epilepsy who retain left hemisphere dominance for language.	**No**	• Children with epilepsy showed reduced whole-network activation in the ventral language system, even when controlling for age and task performance accuracy.
Functional connectivity will be reduced in children with epilepsy.	**No**	• There was no effect of epilepsy on individual function connections in the left hemisphere, nor on specific types of connections; fronto-temporal, fronto-parietal or temporo-parietal connections.
Poorer language in children with epilepsy will	**No**	• Activation strength in the ventral language system was the only

be associated with decreased intrahemispheric functional connectivity.		significant predictor of language in children with epilepsy. • Activation in the ventral language system predicted 47% of variance in language ability in patients.
Chapter 8: Age-related changes in the auditory comprehension network, controlling for task performance difficulty and accuracy		
Age-related changes will be observed in line with the Interactive Specialisation framework of cognitive development: • Age-related increases in activation in regions associated with semantic and syntactic processing, including; bilateral STS, middle and inferior temporal gyri, left ventral occipito-temporal cortex, Pars Triangularis and Orbitalis, superior frontal gyrus, and bilateral angular gyrus. • Age-related decreases will be observed in regions beyond the core language network, associated with executive functioning and working memory.	**Yes (partially)**	• In line with hypotheses, focal regions supporting semantic and syntactic processing showed increasing activation with age; left posterior inferior temporal gyri, bilateral ventral occipito-temporal cortex. • Contrary to hypothesis, increases in activation were observed in regions beyond the core language network, similar to predictions of the maturation framework; dorsolateral prefrontal cortex and STG.

Table 30: Specific hypotheses and findings for each Chapter.

3. PRACTICAL IMPLICATIONS

In this section I discuss practical implications of findings from this thesis in relation to pre-surgical planning, and beyond.

3.1 Pre-surgical planning

3.1.1 Age contributes only minimally to lateralisation of language

I found no change in lateralisation of activation with age in the auditory comprehension network - nor in specific semantic processing regions - when controlling for task difficulty and performance accuracy. This finding challenges the common precept that language becomes increasingly lateralised to the left hemisphere with age. At most, age appears to account for ~10% of individual variability in language lateralisation (Berl et al., 2014; Holland et al., 2007), which is most often reported specifically for frontal regions during verbal fluency tasks (Holland et al., 2007). My findings suggest core language regions are left lateralised from as early as 5 years of age in the majority of typically developing children. Previous reports of increasing leftward lateralisation may reflect development of executive systems involved in controlled semantic retrieval, rather than linguistic processing per se. This finding has important implications for decisions regarding the optimal timing for surgery in children with epilepsy; lateralisation of language in pre-surgical candidates is unlikely to show any major change with age alone.

3.1.2 Candidate 'core' language regions

I provide novel evidence in children for an amodal semantic processing network similar to adults, including bilateral anterior fusiform/parahippocampal gyri, left temporal pole, vOT, Pars Triangularis (BA 45) and left medial temporal cortex (Chapter 5). These regions support semantic processing independent of task and modality, and represent an important target for pre-surgical language mapping beyond classic Broca's and Wernicke's areas. Encouragingly, the new task battery I present here provides comprehensive mapping of these regions. I hypothesise these regions will be critical to post-surgical outcome, as the core

semantic component within the language system. Further research is now required to test this hypothesis, by investigating the effects of structural damage, seizure activity and surgery involving these regions, on language function. For example, voxel-based lesion-symptom mapping may confirm if these regions are critical to language development in children with epilepsy who have structural abnormalities, or who have undergone surgical resection.

3.2 Beyond surgical planning

3.2.1 Providing an ecologically valid model of language development

I have shown the feasibility of investigating naturalistic language using ecologically valid language tasks, which are developmentally appropriate and matched to each individual child based on their own ability level. These tasks allow accurate identification of language networks in children as young as 5 (Chapter 5) and with cognitive impairment (Chapter 6). This shows the feasibility of imaging naturalistic language in populations who may typically struggle to perform language fMRI tasks. The task panel I have developed can therefore be used to investigate language development and impairment across patient populations, including autism spectrum disorder and specific language impairment for example.

3.2.2 The importance of the scanning experience for children

I provide evidence that a poorer experience inside the scanner is associated with increased in-scanner movement. This included how much participants enjoyed the tasks performed during scanning, how difficult they found them, how claustrophobic and how comfortable they felt during scanning, as well as how long they felt they were inside the scanner. It may be possible to reduce scan failures in children and adolescents by improving the physical comfort of participants during scanning (including the use of child-friendly headphones), as well as making fMRI tasks engaging, and keeping scan times short.

3.2.3 The relevance of network topology and synchrony to language ability

There has been an increasing paradigm shift in neuroimaging over the past ten years, away from investigations of network topology (regional activation patterns) towards investigations of network synchrony (using functional connectivity measures). The relative sensitivity of functional connectivity to variance associated with age, epilepsy and language ability has not been investigated previously however. To guide my own analyses of age-relate changes in the language network, I performed the first systematic comparison of fMRI parameters (activation strength versus functional connectivity) to identify whether network topology or synchrony was most sensitive for investigating individual differences in the language system. My findings show network topology changes with age, and predicts language ability in children with epilepsy. In contrast, functional connectivity (as measured by correlational analyses) appears insensitive to between subject variance, and instead emphasises tight synchrony among left hemisphere language regions, which is established early in life. These findings emphasise the value of continued investigation into language network topology. They also highlight the need for more advanced functional connectivity methods, which are more sensitive to variation within strongly synchronised systems. Dynamic Causal Modelling (Friston et al., 2003) enables testing of specific hypotheses regarding network dynamics, and may prove profitable in this instance.

4. LIMITATIONS

4.1 Speech-movement artefact

Speech-movement related artefacts were identified, despite extensive pre-scan practice (Chapter 4) and despite encouraging results from an earlier feasibility study (Chapter 1). This artefact was a result of head movements (particularly rotations) in the direction of slice acquisition (*z* plane). It was defined by signal dropout in either one axial slice (minor) or multiple slices (severe). Despite trialling three artefact repair methods, it was not possible to recover lost signal. These methods repaired

artefacts associated with between-TR movements (between volumes). However, the speech-movement artefact affecting this data set was associated with sub-TR movement (within a volume). Encouragingly, new artefact repair methods are in development, to repair outlying time points on a voxelwise basis (Appendix 12.). To prevent false negatives as a result of this artefact, I excluded all participants with severe artefact from group map analyses.

4.2 Validation of fMRI

As a first step towards validating the Panda Games for the pre-surgical setting, I compared results from the Panda Games with results obtained using a verb generation task, which is validated for use in the pre-surgical setting against invasive methods (Chapter 6). I also presented a case study to demonstrate the sensitivity and specificity of the Panda Games in relation to ICM (Chapter 6, Patient A). Both comparisons have limitations however, which may result in an underestimation of data quality for the Panda Games. All patients recruited for Chapter 6 performed verb generation before the Panda Games, due to clinical priority. Verb generation was also the most recently practised task, immediately before scanning. Temporal and practice effects therefore made the acquisition of high-quality data more challenging for the Panda Games relative to verb generation. Further, patients had to perform a large battery of tasks during scanning, which can also increase scan failure rates (Yerys et al., 2009). The Panda Games have been designed such that a smaller number of tasks can be selected on an individual patient basis for pre-surgical mapping. I therefore hypothesise a lower scan failure rate and improved data quality in the clinical setting. Validation of fMRI with ICM is also problematic; these limitations are discussed in Chapter 6, section 5.5.2.1. Crucially, in my patient case study, ICM was performed using a different task (LN) to fMRI (PN), due to challenging behaviour during both sessions. This highlights the difficulty of obtaining reliable and systematic data across methods in children with severe epilepsy and cognitive impairment, and represents a target for future research.

5. DIRECTIONS FOR FUTURE RESEARCH

5.1 Methodological improvements in the assessment of language in the pre-surgical setting

My findings have highlighted several main areas for future research aimed at improving the accuracy and reliability with which language cortex can be mapped on an individual basis, in the pre-surgical setting.

5.1.1 Systematic validation of fMRI using carefully controlled ICM

Functional MRI has not been well-validated against ICM. Obtaining high-quality data with directly comparable tasks and testing procedures is a challenge in the clinical setting; patients tire easily, often have cognitive or behavioural co-morbidities. Further, it is not always possible to use exactly the same testing materials for both methods; stimulation often needs to be repeated many times over to replicate findings, while fMRI tasks are kept short to minimise time in the scanner. It is important that future studies persist however. The use of automated stimulus presentation methods during ICM – similar to the fMRI procedure – may help standardise the testing procedure.

5.1.2 Individual-level optimisation of fMRI

I have shown several design features for fMRI tasks, which can improve the reliability and accuracy of language localisation on an individual basis. Further work can be done however to develop methods which can determine the optimal task length (during scanning, and on an individual basis) which optimises the sensitivity and specificity of fMRI at a given statistical threshold. Such methods hold the potential to reduce scan failures associated with low signal-to-noise ratio, whilst avoiding unnecessarily long scanning time.

5.1.3 Alternative method for assessing language lateralisation in very young and very impaired children

Surgery performed earlier in life is often associated with better outcome (Loddenkemper et al., 2007; Spencer & Huh, 2008). However, in some cases the

decision for surgery may be delayed, until information regarding the lateralisation or localisation of language can be obtained. In order to extend non-invasive assessments of language lateralisation to younger children (<5 years) and the most severely impaired (e.g. patients with hemispheric conditions such as Rasmussen's encephalitis and Sturge Weber syndrome), future research should aim to trial and validate other methods for assessing language dominance. In their recent review, Potvin and Nussbaum highlight several promising methods, including Functional Transcranial Doppler Imaging, which can be performed when the child is at home, and is moving around to some extent (Potvin & Nussbaum, 2013).

5.2 Novel influences on language lateralisation: A dynamical systems approach

I found no effect of age on lateralisation of activation. Other studies also report little (Berl et al., 2014; Holland et al., 2007) or no (Lidzba et al., 2011) effect of age on language lateralisation. These findings suggest other factors contribute to lateralisation of language. These could be genetically or environmentally driven (Bishop, 2013). Other cognitive systems - through their interaction with language - may also effect language lateralisation; as a form of functional anchor. For example, there is evidence to suggest language lateralisation is linked to the medial temporal memory system (Binder et al., 2010; Hamberger, Seidel, et al., 2007; Liegeois et al., 2004; Weber et al., 2006). Several types of memory are already recognised as important contributors to language, including semantic memory (Hodges et al., 1992; Holland & Lambon Ralph, 2010; Lambon Ralph et al., 2009; Mion et al., 2010; Visser et al., 2010) and working memory (Jacquemot & Scott, 2006; Majerus, 2013; Vigneau et al., 2006). The medial temporal lobe memory system (classically associated with episodic memory) has received less attention with respect to language however. This is changing (Kurczek et al., 2013; Opitz & Friederici, 2004) as the role of memory in naturalistic language is emphasised by some models (Hagoort, 2013). The importance of network interactions has been acknowledged for a long time in cognitive neuroscience (Mesulam, 1990). Despite this, the effect of system interactions on measures of language lateralisation has not been investigated. The evidence listed above suggests interactions between core memory and language

networks may influence lateralisation of function. Future investigations into the effect of system interactions on language lateralisation may highlight biomarkers of the potential for re-organisation of function, and may even provide targets for interventions which aim to influence language lateralisation (to induce reorganisation of function for example).

5.3 Building an accurate and testable model of language development

The last 20 years of neuroimaging in adults has produced a comprehensive mapping of structure-function relationships in the language system (Price, 2012). These are "*…central to the cognitive neuroscience endeavour and require careful definition*" (Price and Friston, 2007; page 272). Future research is required to replicate this experimental work in paediatric populations, to understand how and when these relationships become established. These structure-function mappings (which may be many-to-many, as opposed to one-to-one), will provide a more comprehensive model of language development which can be used to inform and test specific hypotheses regarding typical and atypical trajectories.

5.4 A shift in focus: From 'Broca-Wernicke' to 'multiple routes to meaning'

Picture naming can be performed without engagement of Broca's area (Etard et al., 2000). I provide evidence that sentence generation can also be performed without engagement of Broca's area (Chapter 5). This finding suggests there may be multiple routes to sentence generation, as for reading (Richardson et al., 2011), some of which bypass Broca's area. Specifically, I identified a ventral route to meaning, extending from the posterior vOT, along the fusiform gyrus, to the temporal pole. This pathway may provide an automatic route for sentence generation, which relies on a direct mapping between familiar visual patterns (including pictorial depictions of actions) and word forms (Price & Devlin, 2011). Meta-analyses provide support for the existence of multiple pathways within the ventral and dorsal language systems, and even identify a ventral semantic pathway also involving the vOT and temporal poles (Figure 63) (Vigneau et al., 2006). The authors propose this ventral

pathway reflects the very foundations of language, as the route by which meaning is achieved (Vigneau et al., 2006). They call for a revision of the traditional Broca-Wernicke model.

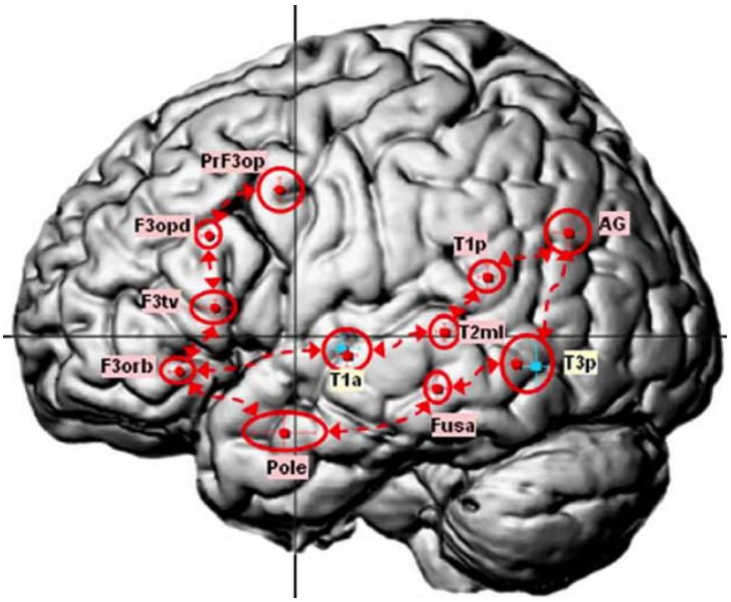


Figure 63: Multiple routes to meaning. Pathways supporting semantic processing, identified across 129 neuroimaging studies of language (Vigneau et al., 2006). I identified a ventral pathway supporting semantic processing in children, similar to the pathway shown here including: T3p (posterior inferior temporal gyrus), fusiform gyrus (Fusa), temporal pole (Pole) and Pars Orbitalis (F3orb).

I therefore propose a shift of focus for developmental investigations of language, away from an exclusive focus on the maturation of Broca's and Wernicke's area, towards an understanding of how different semantic processing pathways change with age. The influence of bottom-up (feed forward) and top-down (feedback) connections on these pathways can be formally tested using Dynamic Causal Modelling (Friston et al., 2003). Developmental studies may be particularly well placed to investigate the effects of learning on these pathways, as my findings suggest nodes within this system (specifically vOT) show prolonged maturation. This new 'multiple routes' approach may be particularly valuable for understanding the

effects of epilepsy and surgical intervention on language outcome, and may also provide insights into mechanisms of functional reorganisation.

FINAL CONCLUSIONS

I provide evidence for a ventral pathway – previously unidentified in children – which extends from vOT, along the fusiform gyrus, to the temporal poles and anterior IFG. This ventral pathway supports semantic processing regardless of task and modality, and may reflect the very foundations of language. I have developed a new fMRI task battery which enables reliable and comprehensive mapping of this system in the pre-surgical setting. Based on novel network analyses, I also suggest the ventral language system – not the dorsal system – explains variability in language function in children with epilepsy. The relevance of this system to post-surgical outcome is not yet well understood. It is possible that structural abnormalities, seizure activity and surgical resections involving this pathway hold the potential to alter access to meaning, and therefore impair language or induce functional reorganisation in children with drug-resistant epilepsy. Encouragingly, I provide evidence for multiple routes to meaning, which may provide a degree of flexible compensation within the system. The ventral occipito-temporal region – a previously underappreciated node in the language system - is particularly important for accessing meaning in children, and shows protracted development. Based on these findings I propose a revision of the classic Broca-Wernicke model of language is required in the developmental context. I propose a new model of language development should acknowledge multiple routes within ventral and dorsal language systems, and the contribution of regions located much more ventrally than previously appreciated, to semantic processing.

GLOSSARY

Here I provide an alphabetised list of abbreviations used within this thesis.

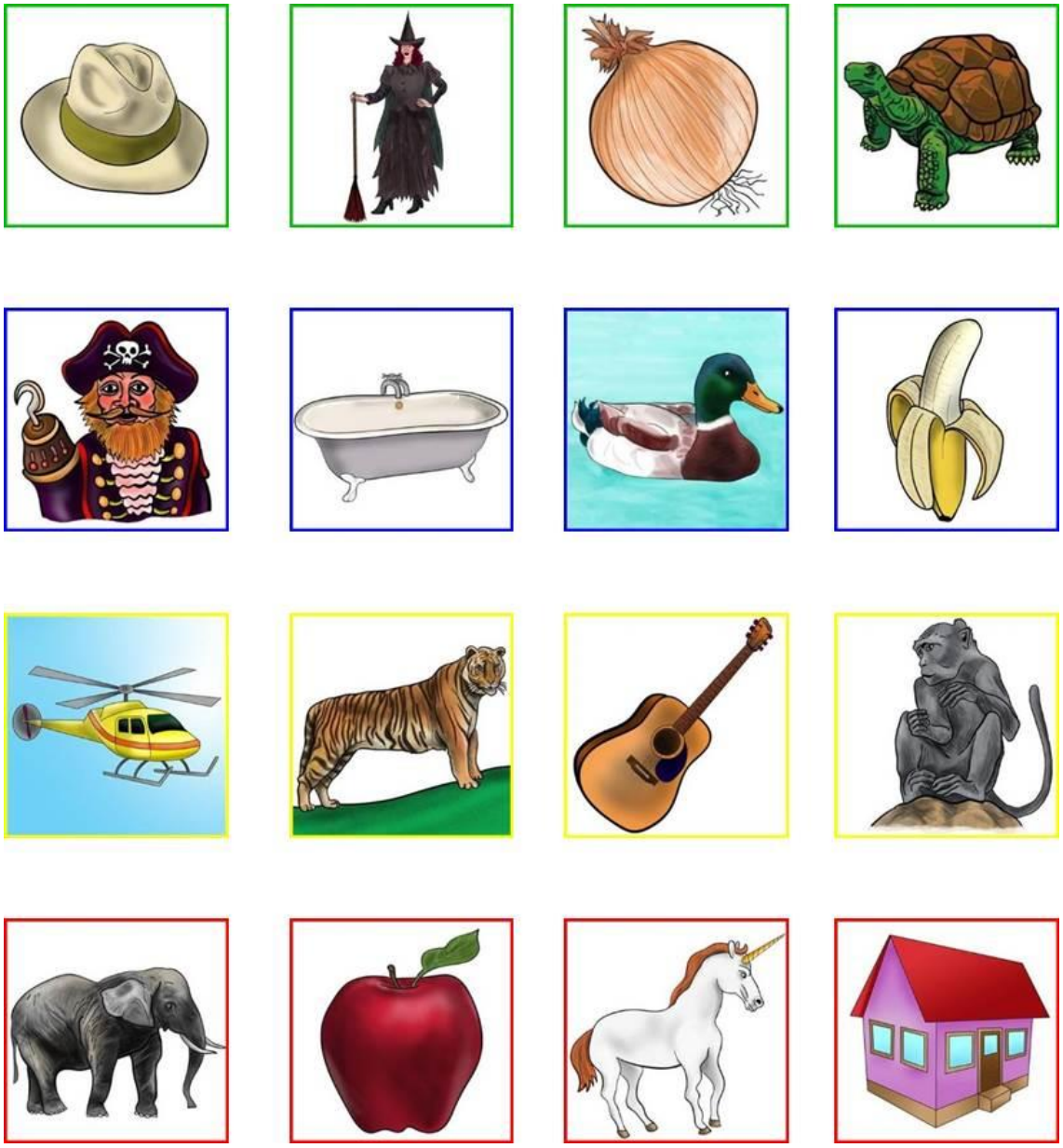
a	Anterior (e.g. aSTG)
AF	Arcuate Fasciculus
AG	Alien Game
AG	Angular gyrus
AoA	Age of acquisition
CNs	Colour Naming (with sentences)
CNw	Colour Naming (with words)
DNK	Do not know' response
DTI	Diffusion tensor imaging
EmC	Extreme capsule
FC	Functional connectivity
FF	Face Finder
fMRI	Functional magnetic resonance imaging
Freq	Word frequency
H	The H statistic (a measure of response uncertainty)
ICM	Intracranial mapping
IFG	Inferior frontal gyrus
IFOF	Inferior frontal occipital fasciculus
IPL	Inferior parietal lobe
ITG	Inferior temporal gyrus
LI	Laterality index
LN	Listen and Name
m	Middle (e.g. mSTG)
mm	Millimetres
MFG	Middle frontal gyrus
MISS	Missed response (omission)
MLF	Middle longitudinal fissue
MRI	Magnetic resonance imaging

MTG	Middle temporal gyrus
NA	Name agreement
OT	Occipito-temporal cortex
p	Posterior (e.g. pSTG)
PCA	Principal component analysis
PD	Picture Describing
PN	Picture Naming
RA	Response agreement
RN	Read and Name
RT	Reaction time
SS	Standard Scores
SFG	Superior frontal gyrus
SLF	Superior Longitudinal Fasciculus
SMG	Supramarginal gyrus
spm	Statistical parametric mapping
STG	Superior temporal gyrus
TP	Temporal pole
vOT	Ventral occipito-temporal cortex

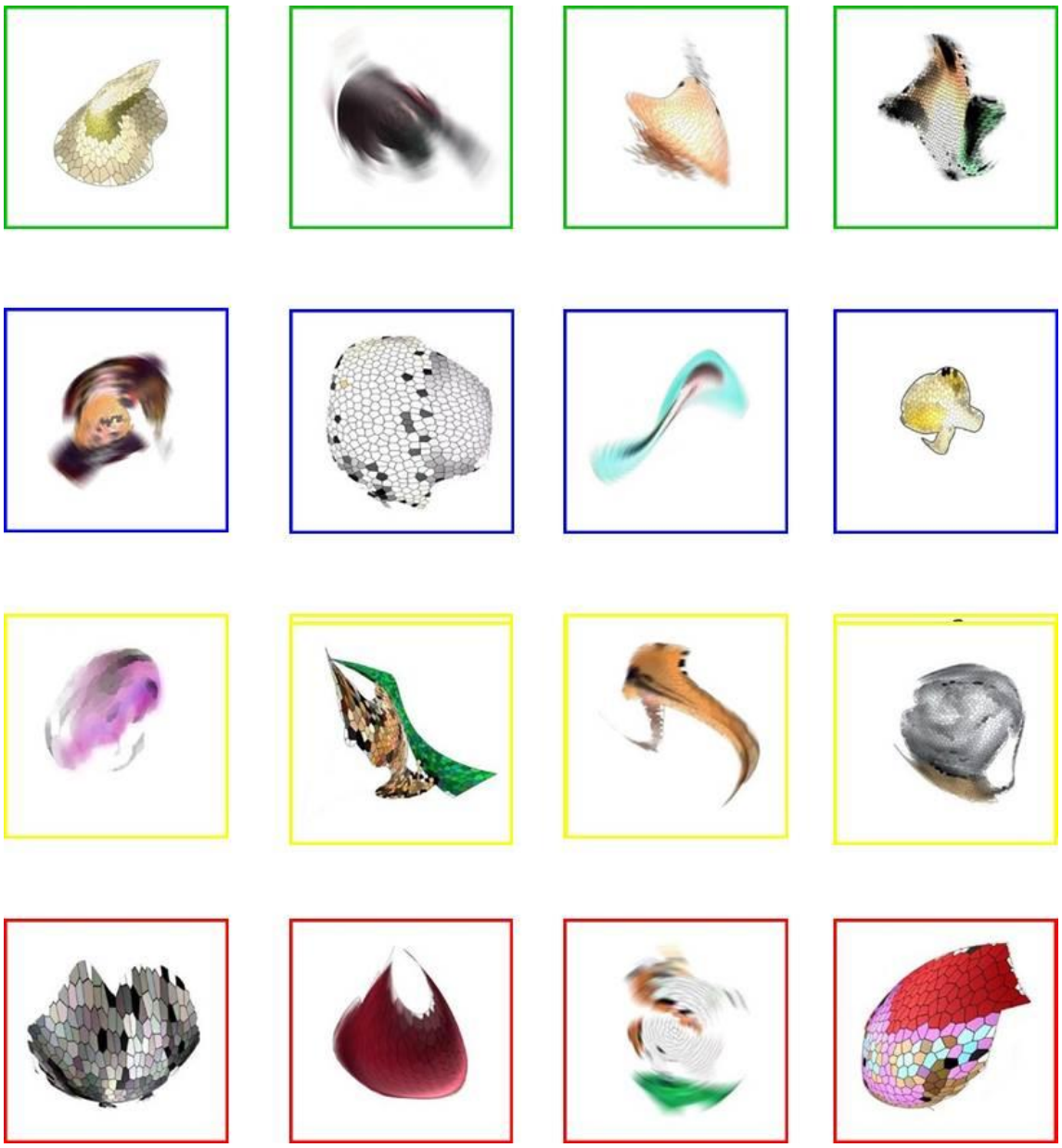
APPENDIX

1. PANDA GAMES STIMULI EXAMPLES

1.1 Picture Naming (PN)



1.2 Colour Naming with words (CNw)



1.3 Picture Describing (PD)





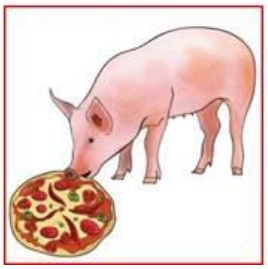




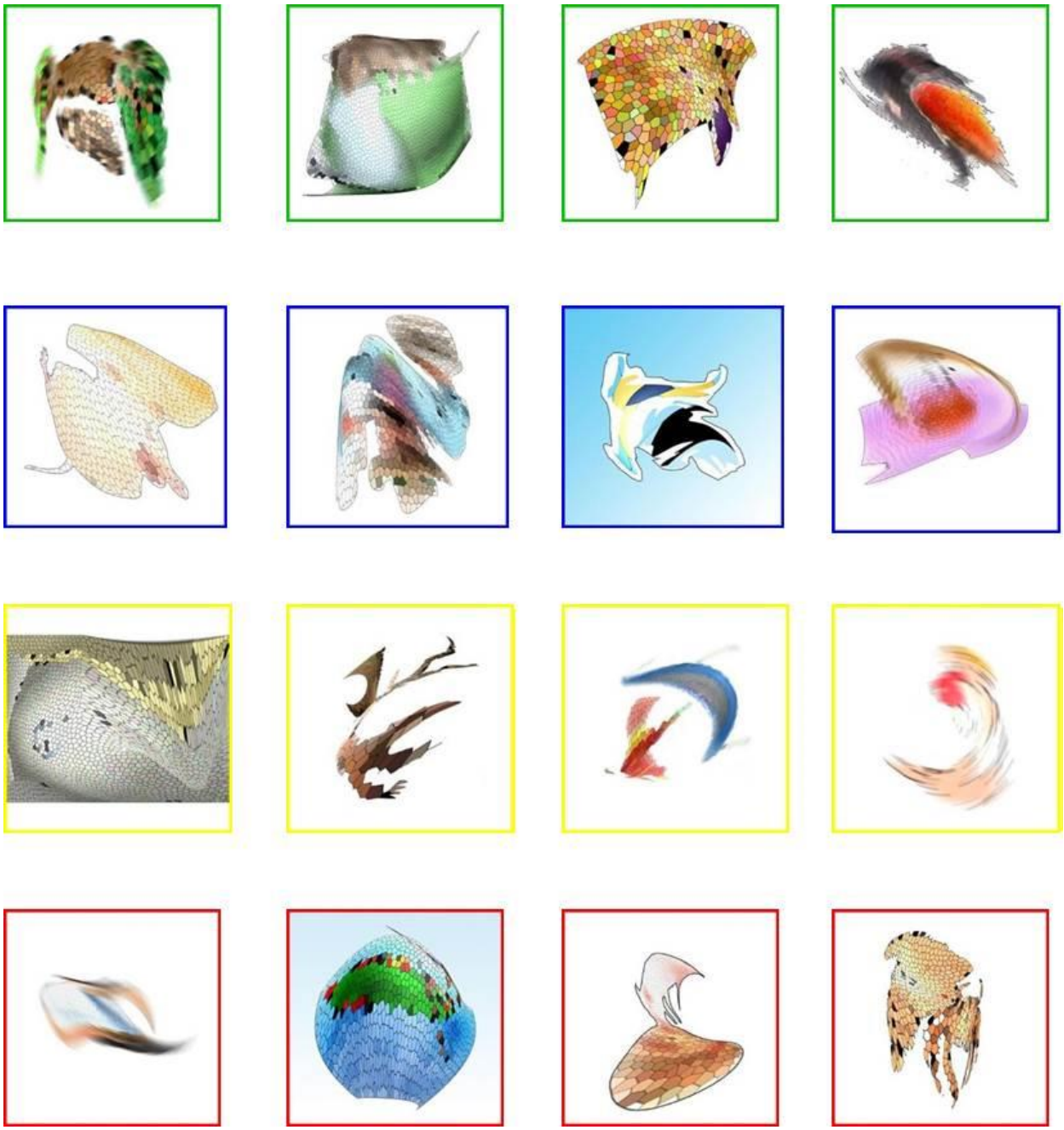




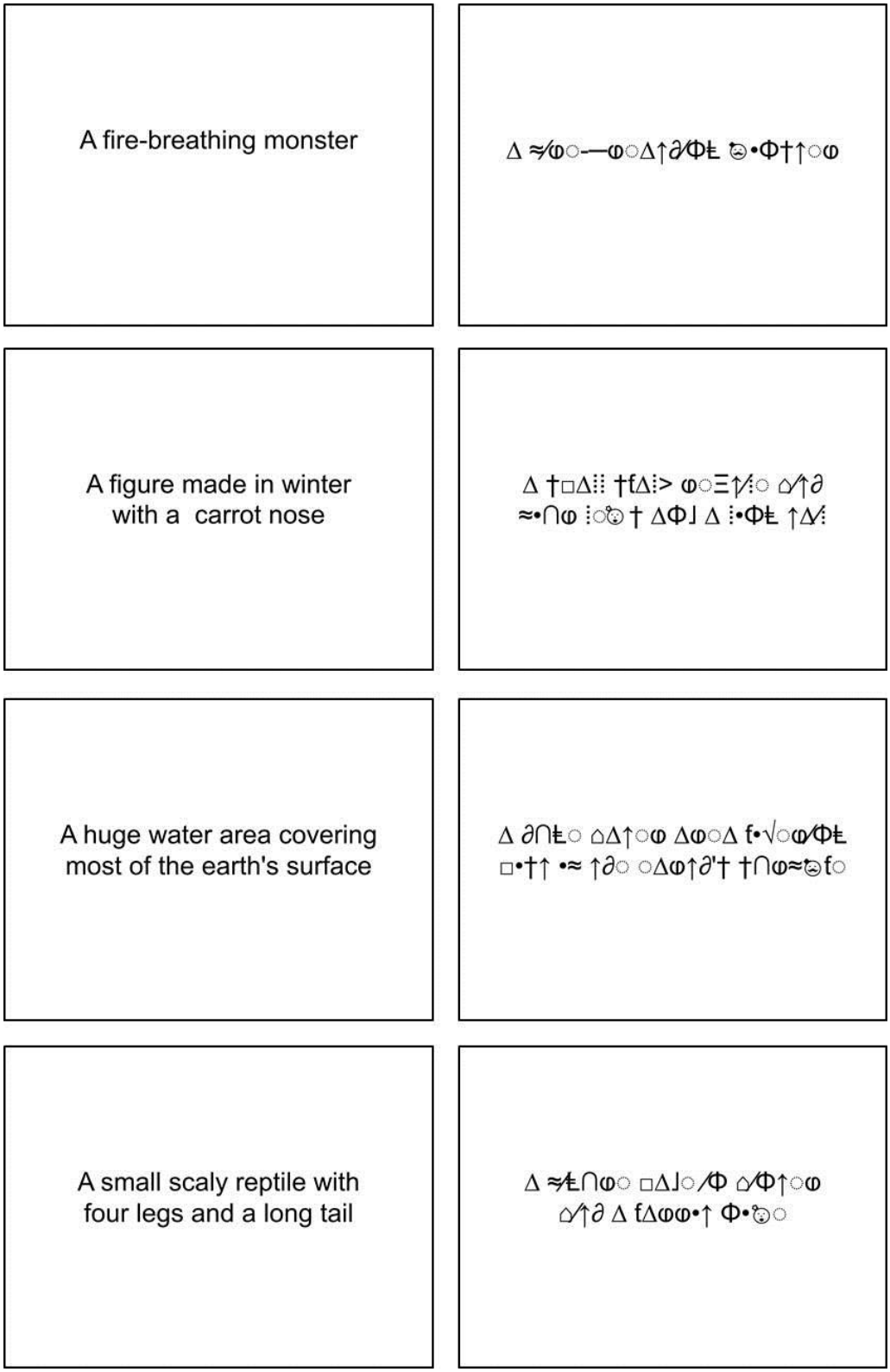





1.4 Colour Naming with sentences (CNs)



1.5 Read and Name (RN) and Face Finder (FF)



2. BEHAVIOURAL PILOT TASK INSTRUCTIONS

2.1 Picture Naming (PN)

In this task, you will see a picture of an object or animal on the computer screen. We would like you to say what the object or animal is called, as quickly as you can. For example, if you see this picture (below), you could say… 'squirrel'

Please say only the name of the animal or object. For example…

Please do say:

Squirrel

Please don't say:

It's a squirrel

I can see a squirrel

Ummmm, squirrel

Frequently asked questions:

- **What if my answer is different to what you think I will say?**

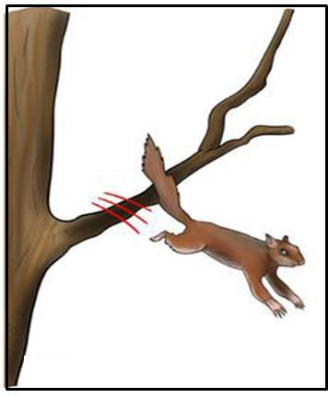
  For example, if I say 'mouse', not 'squirrel'.

  That's ok! We don't expect everyone to say the same thing. Just say what you think the object or animal is called.
- **What if I don't know, or can't remember what the object or animal is called?**

  That's ok too. Just tell the researcher 'I don't know' or 'I forgot'.
- **What if I can't say the name quickly?**
- That's fine! Don't worry about the time, just think about its name.

2.2 Picture Describing (PD)

In this task you will see a picture. In this picture one of the objects or animals will be doing something. **We would like you to tell us what is happening in the picture in one simple sentence.** In every picture the animal or object will be doing 1 of 4 different actions: Eating, drinking, jumping or falling. So if you see this picture…



You could say: '**The squirrel is jumping from the tree'**

Please try to say the sentences in this way: '**The squirrel is jumping from the tree'**

Rather than:

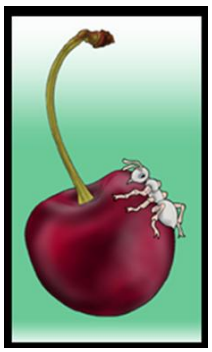
The squirrel is jumping

The squirrel jumps from the tree

Jumping

Some pictures might look like the animal or object is falling **or** jumping. This is ok. Just choose the word you think is best for the picture. So you could also say: '**The squirrel is falling from the tree**.'

In some pictures, it might look like the animal is doing something different from eating, drinking, jumping or falling. For example, this picture might look like 'The ant is crawling on the cherry'



Try to remember there are only **4 actions**: Eating, drinking, jumping and falling. So for this picture you could say '**The ant is <u>eating</u> the cherry'**.

If there is a picture of an animal drinking, please say what it is <u>drinking from</u> (not the thing it is drinking). For example, if you saw this picture…

You could say: 'The dog is drinking <u>from the toilet</u>.'
Rather than: 'The dog is drinking toilet water'

<u>Frequently asked questions:</u>

- **Wow, that was a lot of instructions! I am not sure I can remember everything…**
  That's fine! We understand there's a lot to remember. There will be a break half way through the game, when you can ask the researcher to remind you what to do.
- **What if I'm not sure what the animals or objects in the picture are?**
  That's ok! Say what you think it is and try to make the sentence. If you really don't know, tell the researcher 'I don't know'.
- **What if I'm not sure what action is happening?**
  That's ok too. Remember there are only 4 actions (eating, drinking, falling, jumping), so try to choose the best one from these 4 to make a sentence.

2.3 Read and Name (RN)

In this game, you will see a sentence on the computer screen which describes an animal or object. We would like you to read the sentence and guess what it is describing as quickly as you can. For example, if you see this sentence…

A small red or grey animal with a very bushy tail that likes to eat nuts

…you could say **'squirrel'**

Please say only the name of the animal or object. For example…

Please do say:

Squirrel

Please don't say:

It's a squirrel

I can see a squirrel

Ummmm, squirrel

Frequently asked questions:

- **What if I can't read the sentence, or a word in the sentence?**
  That's ok! You will have 4 practice trials at the beginning of the game. If you find these difficult, tell the researcher and they can read the rest of the sentences out loud for you, to make it easier.
- **What if I don't know what the sentence is describing?**
  That's ok too! Just say 'I don't know'

Remember, these games are designed for older people too, so it is likely you will find some of the sentences difficult!

3. STIMULI CHARACTERISTICS AND LEVEL ASSIGNMENT:

Stimuli characteristics are shown for Picture Naming, Picture Describing and Read/Listen and Name stimuli from the behavioural pilot study presented in Chapter 2. Measures reported include; 1) the dominant (modal) response from a sample of 20 children aged 5-15 years (used as the 'correct' response for the Panda Games), 2) reaction times (in milliseconds) for all 20 children and also 9 adults (aged 23-55 years), 3) the length of the modal response (in letters), 4) Name Agreement (NA), 5) the H statistic (a measure of naming uncertainty), 6) word frequency (based on the British National Corpus), 7) age of acquisition (AoA), 8) mean difficulty rating (according to NA, H and freq), and 9) which level the stimulus was assigned to in the Panda Games.

Age of acquisition data were taken from a norming study performed in British adults (Johnson et al., 2001), except where data are reported in italics (taken from Stadthagen-Gonzalez & Davies et al,. 2006) and where underlined (taken from Gilhooly and Logie, 1980). Age of acquisition data reflects a likert scale, such that: 1=acquired by 0-2 years, 2=3-4 years, 3=5-6 years, 4=7-8 years, 5=9-10 years, 6=11-12 years, and 7=13 years plus.

Spare stimuli (S) were used as practice stimuli for the Panda Games. Excluded stimuli (E) are those which were too difficult for children (as discussed in Chapter 2) and which were excluded from use in the Panda Games task battery (although some of these items were used for task practice).

For Picture Describing I report Response Agreement (RA) as opposed to Name Agreement (NA).

3.1 Picture Naming (PN)

Item no.	Dominant response (mode)	Child RT (ms)	Adult RT (ms)	Response length (letters)	NA	H	Freq	AoA (likert)	Mean difficulty rating	LEVEL
1	apple	972	918	5	100%	0.00	2610	2.00	1.33	1
2	baby	873	833	4	100%	0.00	8605	2.00	1.00	S
3	bag	1143	873	3	95%	4.40	5089	3.00	2.33	S
4	bamboo	1099	1277	6	75%	16.70	287	*3.00*	4.00	4
5	banana	892	697	6	100%	0.00	523	2.00	1.67	1
6	barrel	1446	784	6	90%	6.80	860	4.00	3.33	S
7	basket	1029	1023	6	100%	0.00	1257	3.00	1.33	S
8	bath	1068	793	4	70%	10.90	3788	2.00	3.33	S
9	bear	1096	851	4	100%	0.00	5597	3.00	1.00	1
10	bed	963	774	3	100%	0.00	14695	2.00	1.00	1
11	**cup**	1755	1211	3	61%	20.00	11965	2.00	4.00	S
12	bird	1375	1439	4	90%	3.47	3876	2.00	2.00	S
13	biscuit	1422	954	7	100%	0.00	555	2.00	1.67	S
14	boat	995	1038	4	95%	4.40	5266	1.00	2.33	S
15	book	980	795	4	100%	0.00	25833	2.00	1.00	1
16	bottle	1197	933	6	85%	2.97	3931	1.00	2.33	2
17	bowl	975	1083	4	95%	4.40	2363	2.00	2.67	S
18	boy	1028	1022	3	95%	4.40	12699	1.00	2.33	S
19	bread	1225	1041	5	80%	7.38	3635	2.00	3.33	4
20	bus	932	957	3	95%	4.40	5325	2.00	2.33	2
21	cake	1063	788	4	100%	0.00	2720	2.00	1.33	1
22	camel	1072	821	5	100%	0.00	360	3.00	2.00	2
23	camera	1071	892	6	100%	0.00	2700	3.00	1.33	1
24	car	836	790	3	100%	0.00	26886	2.00	1.00	1
25	carrot	885	736	6	100%	0.00	377	2.00	2.00	2
26	cat	995	855	3	95%	4.40	3853	1.00	2.67	S
27	chair	930	761	5	100%	0.00	7388	2.00	1.00	S
28	cheese	1005	779	6	100%	0.00	2516	2.00	1.33	1
29	cherry	1200	935	6	75%	11.80	868	3.00	3.67	4
30	clock	1157	820	5	85%	7.71	2801	2.00	3.33	4
31	clown	1073	727	5	95%	4.40	361	2.00	3.33	S
32	coconut	1124	953	7	94%	0.08	318	-	3.00	3
33	cow	1111	921	3	100%	0.00	1351	2.00	1.33	S
34	cowboy	1469	1094	6	68%	15.96	319	3.00	4.00	4
35	crab	1074	840	4	95%	0.07	323	3.00	3.00	3

36	cup	2003	1049	3	55%	24.21	11965	2.00	3.00	E
37	deer	1386	1426	4	86%	7.90	909	3.00	3.33	E
38	dog	1056	870	3	100%	0.00	7872	2.00	1.00	1
39	dolphin	971	853	7	95%	4.40	772	3.00	3.00	S
40	donkey	1147	1266	6	90%	3.47	511	3.00	3.00	3
41	dragon	1330	827	6	95%	4.40	871	3.00	3.00	3
42	duck	979	909	4	100%	0.00	1351	2.00	1.33	1
43	elephant	995	878	8	95%	4.40	901	2.00	3.00	3
44	fence	1200	915	5	79%	2.59	1668	3.00	2.67	3
45	fire	842	816	4	100%	0.00	13575	2.00	1.00	1
46	fish	1033	1010	4	100%	0.00	10750	2.00	1.00	1
47	fountain	1222	741	8	58%	0.79	667	4.00	3.00	E
48	fox	1078	933	3	90%	8.80	2282	3.00	3.00	3
49	fridge	1259	839	6	100%	0.00	895	3.00	1.67	1
50	frog	1001	656	4	95%	4.40	492	2.00	3.33	S
51	gate	1139	902	4	85%	2.97	3427	3.00	2.67	S
52	giraffe	1043	998	7	100%	0.00	71	3.00	2.00	2
53	girl	1052	978	4	85%	7.88	14569	2.00	3.00	3
54	glass	1105	877	5	84%	2.91	9348	2.00	2.33	2
55	gloves	985	712	6	100%	0.00	957	2.00	1.67	1
56	goat	1270	1104	4	89%	3.41	608	3.00	2.67	3
57	grapes	952	845	6	100%	0.00	517	3.00	2.00	2
58	guitar	898	799	6	100%	0.00	2756	3.00	1.33	1
59	hand	944	735	4	95%	4.40	33437	2.00	2.33	2
60	hat	1121	705	3	100%	0.00	2903	2.00	1.33	1
61	hedge	1349	1318	5	67%	6.60	945	3.00	3.67	4
62	helicopter	933	847	10	95%	4.40	1098	3.00	3.00	3
63	hoop	1158	1066	5	80%	7.38	128	-	4.00	4
64	horse	1053	834	5	100%	0.00	7503	2.00	1.00	1
65	house	939	769	5	95%	4.40	49424	-	2.33	2
66	**coat**	1162	932	4	37%	18.10	3294	2.00	3.33	E
67	jug	1315	999	3	80%	17.61	517	3.00	4.00	S
68	kangaroo	1180	952	8	100%	0.00	114	4.00	2.00	2
69	kettle	1139	974	6	55%	20.15	847	3.00	3.67	E
70	king	1116	773	4	95%	4.40	15793	3.00	2.33	2
71	koala	1163	970	5	63%	6.36	17	-	4.00	E
72	ladder	1201	941	6	95%	4.40	1271	3.00	2.67	3
73	lamb	1117	882	4	75%	2.42	1635	2.00	2.67	3
74	leaves	1284	1029	6	65%	6.68	6354	2.00	3.00	3
75	lion	926	736	4	95%	4.40	1226	3.00	3.00	3
76	lizard	1224	1077	6	95%	4.33	210	4.00	3.33	4
77	log	951	1024	3	90%	8.80	1231	3.00	3.33	S
78	man	1207	1025	3	89%	8.66	58821	2.00	2.67	S
79	mermaid	1117	822	7	100%	0.00	89	3.00	2.00	2
80	monkey	1218	966	6	85%	7.88	556	3.00	3.67	4
81	mouse	1060	984	5	85%	7.88	1854	2.00	3.33	4
82	**cup**	1286	1060	3	55%	2.01	11965	2.00	2.33	E
83	nest	1170	1010	4	95%	4.40	1426	3.00	2.67	3
84	nun	1238	722	3	63%	15.09	418	-	4.00	E

85	nurse	1309	860	5	55%	11.02	3259	3.00	3.33	E
86	onion	1033	790	5	100%	0.00	675	3.00	1.67	1
87	owl	898	731	3	95%	4.40	1137	3.00	3.00	3
88	panda	1086	835	5	85%	13.20	176	3.00	4.00	S
89	parrot	972	748	6	90%	3.47	387	3.00	3.00	3
90	pear	1103	965	4	95%	4.40	224	3.00	3.33	4
91	penguin	937	788	7	100%	0.00	507	3.00	2.00	2
92	piano	1329	855	5	100%	0.00	1916	3.00	1.33	S
93	pig	842	783	3	100%	0.00	1320	2.00	1.33	1
94	**policeman**	1201	1499	9	45%	23.76	2031	-	3.33	E
95	pirate	1107	1029	6	90%	8.80	270	3.00	3.67	4
96	pizza	1005	680	5	100%	0.00	480	3.00	2.00	2
97	plane	894	914	5	55%	6.51	3391	3.00	3.33	E
98	plate	1022	840	5	100%	0.00	3965	2.00	1.00	1
99	potato	1030	1088	6	95%	4.40	865	2.00	3.00	3
100	pumpkin	992	992	7	100%	0.00	73	4.00	2.00	2
101	**dog**	1024	1043	3	60%	6.57	7872	2.00	3.00	3
102	rabbit	1023	1108	6	90%	3.47	1403	2.00	2.33	S
103	**rock**	1125	1054	4	45%	10.95	6670	3.00	3.00	E
104	sack	1185	930	4	44%	15.05	1026	*3.00*	3.67	E
105	**pan**	1117	857	3	74%	11.18	323	-	4.00	4
106	sea	1238	967	3	55%	5.18	12590	-	3.00	E
107	seal	1168	906	4	82%	0.28	1383	4.00	2.67	3
108	sheep	1088	1394	5	70%	6.84	2983	2.00	3.33	S
109	shelf	1399	1159	5	53%	0.93	1381	*3.00*	2.67	E
110	shell	1045	829	5	100%	0.00	2182	3.00	1.33	1
111	shoe	915	788	4	95%	4.40	1126	2.00	3.00	3
112	snake	919	830	5	100%	0.00	730	3.00	1.67	S
113	snowman	935	846	7	95%	4.40	76	2.00	3.33	4
114	sofa	1062	921	4	95%	4.33	921	2.00	3.00	3
115	soldier	1299	1206	7	39%	22.39	1725	3.00	3.33	E
116	chair	1271	844	5	65%	6.68	815	3.00	3.67	S
117	**oven**	1366	1229	4	47%	9.65	1311	3.00	3.33	E
118	suitcase	1189	973	8	70%	6.84	522	3.00	3.67	4
119	swan	1680	951	4	80%	7.38	982	3.00	3.67	4
120	swimming pool	1088	857	12	65%	2.14	799	3.00	3.00	3
121	swing	1183	914	5	100%	0.00	1896	2.00	1.33	1
122	table	1124	986	5	95%	4.40	19289	2.00	2.33	2
123	tap	1174	773	3	100%	0.00	2042	3.00	1.33	1
124	teapot	1240	707	6	65%	20.65	241	3.00	4.00	4
125	tent	1005	823	4	95%	4.40	1097	3.00	3.00	3
126	tiger	997	893	5	100%	0.00	888	3.00	1.67	S
127	toilet	861	898	6	100%	0.00	1540	2.00	2.00	S
128	tomato	1145	796	6	100%	0.00	720	3.00	2.33	2
129	train	1024	903	5	95%	4.40	7861	2.00	2.33	2
130	trampoline	954	796	10	100%	0.00	46	-	2.00	2
131	tree	886	740	4	95%	4.40	6136	2.00	2.33	2
132	**car**	1226	1134	3	50%	6.47	26886	2.00	3.00	E
133	**tortoise**	835	872	8	65%	2.14	154	3.00	3.33	4

134	unicorn	977	836	7	90%	8.80	94	4.00	3.67	4
135	wall	1656	1363	4	42%	18.91	11298	2.00	3.00	E
136	waterfall	1271	1008	9	79%	17.34	387	-	4.00	4
137	whale	1150	1290	5	95%	4.40	571	3.00	3.00	S
138	**glass**	1151	1144	5	60%	5.47	9348	2.00	3.00	4
139	witch	922	785	5	100%	0.00	573	3.00	1.67	2
140	zebra	870	729	5	100%	0.00	222	3.00	2.00	2

3.2 Read/Listen and Name (RN and LN)

Item no.	Stimulus	Dictio-nary age	Dominant response (mode)	Child RT (ms)	Adult RT (ms)	NA	H	Freq	AoA (likert)	LEVEL (RN)	LEVEL (LN)
1	A red or green fruit with a core	5-7yrs	apple	4006	2359	85%	13	2610	2.00	2	S
2	A very young child	5-7yrs	baby	4985	2368	65%	7	8605	2.00	2	S
3	An item you wear to carry things in	5-7yrs	bag	4324	2638	78%	12	5089	3.00	2	2
4	A plant with hard hollow stems which Panda's eat	11-16yrs	bamboo	4825	2833	84%	8	287	*3.00*	3	3 and 4
5	A long curved fruit with thick yellow skin	5-7yrs	banana	3326	2520	72%	11	523	2.00	S	3
6	A large container for liquids with curved sides and flat ends	7-11yrs	barrel/bowl/jug	3338	7802	20%	20	860	4 or 2 or 3	E	E
7	A container made from wicker	7-11yrs	basket	4708	2006	64%	15	1257	3.00	E	E
8	A large container of water in which to wash your whole body	11-16yrs	bath	4376	2077	85%	13	3788	2.00	2	1 and 2
9	A big heavy animal with thick fur and sharp claws	5-7yrs	lion	3189	3628	39%	14	1226	3.00	E	E
10	A piece of furniture to sleep on	5-7yrs	bed	3025	1563	100%	0	14695	2.00	1	1
11	A tall and heavy silver drinking mug	11-16yrs	cup	3417	3297	50%	13	11965	2.00	E	E
12	An animal that has wings, feathers and a beak.	5-7yrs	bird	4497	3376	85%	8	10	2.00	3	S
13	A small flat kind of cake that has been baked until crisp	11-16yrs	biscuit	3005	3381	31%	22	3876	2.00	E	E
14	A vehicle for travel on water	7-11yrs	boat	3762	2137	94%	4	555	2.00	1	1
15	Something with a cover and pages inside	5-7yrs	book	5600	2372	89%	9	5266	1.00	1	S
16	A container a	5-	bottle	3489	2006	85%	13	3931	1.00	1	1

	baby drinks from	7yrs									
17	A deep round dish for eating soup from	7-11yrs	bowl	4583	1819	100%	0	2363	2.00	S	1
18	A male child	5-7yrs	boy	4609	2828	82%	13	12699	1.00	2	2
19	A loaf-shaped food made of wheat	7-11yrs	bread	3265	3478	83%	13	3635	2.00	3	2 and 3
20	A big vehicle that can carry a lot of people to and from places	5-7yrs	bus	3809	2528	68%	26	5325	2.00	3	3
21	A baked treat often eaten at birthdays and Christmas	5-7yrs	cake	4060	2119	89%	9	2720	2.00	1	1
22	A humped desert animal	11-16yrs	camel	4292	2000	100%	0	360	3.00	2	1 and 2
23	Something to take photos with	5-7yrs	camera	3507	1595	100%	0	2700	3.00	1	S
24	Another name for automobile	11-16yrs	car	3103	1599	64%	20	26886	2.00	S	3
25	A long orange vegetable	7-11yrs	carrot	3872	2178	83%	8	377	2.00	3	3
26	A grown-up kitten	5-7yrs	cat	5533	1092	95%	4	3853	1.00	1	1
27	A piece of furniture to sit on with four legs	5-7yrs	chair	5684	2992	85%	3	7388	2.00	1	S
28	A white or yellow food made from milk, it can be hard or soft	7-11yrs	cheese	4316	2902	62%	15	2516	2.00	E	E
29	A small red fruit with a stone from a blossoming tree	7-11yrs	cherry	4426	3089	50%	18	868	3.00	4	4
30	An object on the wall that tells the time	5-7yrs	clock	5353	2091	100%	0	2801	2.00	1	1
31	A funny circus person	7-11yrs	clown	4312	1244	100%	0	361	2.00	S	1 and 2
32	A large nut which grows on palm trees	7-11yrs	coconut	3189	2534	78%	17	318	-	4	4
33	An animal that gives milk	5-7yrs	cow	5392	1591	100%	0	1351	2.00	1	S
34	A man who rides a horse and looks after cattle on a ranch	7-11yrs	cowboy	5569	2908	55%	10	319	3.00	E	E
35	A shellfish with pincers and eight legs	11-16yrs	crab	4055	1899	70%	11	323	3.00	4	S
36	A container you drink tea from	5-7yrs	cup	3511	2808	59%	15	11965	2.00	S	3.00
37	An animal with antlers	7-11yr	deer	4385	1875	44%	25	909	3.00	E	E

		s									
38	A barking pet	5-7yrs	dog	4969	1353	95%	4	7872	2.00	1	1
39	A fun-loving sea mammal with flippers	7-11yrs	seal	4360	2652	38%	17	2983	2.00	E	E
40	A horse-like animal often ridden at the seaside	5-7yrs	donkey	5551	2721	76%	17	511	3.00	S	4
41	A fire-breathing monster	5-7yrs	dragon	5157	1725	100%	0	871	3.00	1	1
42	An animal that quacks	5-7yrs	duck	4856	1328	100%	0	1351	2.00	1	1
43	A big grey animal with tusks and a trunk	7-11yrs	elephant	5336	2317	95%	4	901	2.00	2	1 and 2
44	A wooden barrier surrounding a garden	7-11yrs	fence	6828	2302	73%	11	1668	3.00	3	3
45	Something hot made of flames	5-7yrs	fire	5821	2145	95%	4	13575	2.00	1	S
46	An animal with fins and gills living in water	11-16yrs	fish	4458	2914	72%	11	10750	2.00	2	2
47	A garden or park feature shooting out water	7-11yrs	fountain	6368	2831	47%	23	667	4.00	E	E
48	Like a dog with reddish fur and a bushy tail	11-16yrs	fox	3443	2761	56%	15	2282	3.00	S	4
49	A metal cupboard that keeps food cold	5-7yrs	fridge	4411	2364	89%	3	895	3.00	2	1 and 2
50	A green animal that croaks	5-7yrs	frog	6169	1849	100%	0	492	2.00	2	S
51	A hinged door in a fence	7-11yrs	gate	5585	1933	85%	8	3427	3.00	E	E
52	An African animal with a very long neck	7-11yrs	giraffe	3445	2139	95%	4	71	3.00	2	2
53	A female child	5-7yrs	girl	5173	1646	76%	12	14569	2.00	2	2
54	Something clear or tranSnt you drink out of	11-16yrs	glass	4270	2309	50%	14	9348	2.00	3	3 and 4
55	Clothing to keep hands warm	5-7yrs	gloves	5993	1517	95%	4	957	2.00	2	2
56	A milk-producing animal with a beard	5-7yrs	goat	4947	3818	67%	11	608	3.00	4	S
57	A fruit which grows in bunches on vines	7-11yrs	grapes	3745	2293	69%	16	517	3.00	4	4
58	A musical instrument played by strumming its strings	7-11yrs	guitar	4846	2330	89%	9	2756	3.00	2	2

59	A body part with a thumb, fingers and palm	5-7yrs	hand	4140	1993	100%	0	33437	2.00	S	1
60	Clothing you wear on your head	5-7yrs	hat	5207	1523	80%	18	2903	2.00	3	S
61	A kind of wall made by bushes growing close together	5-7yrs	hedge	7919	4187	55%	16	945	3.00	E	E
62	A small aircraft with a propeller instead of wings	7-11yrs	helicopter	5150	4702	59%	13	1098	3.00	4	4
63	A ring that performing animals jump through	7-11yrs	hoop	5294	3961	41%	27	128	-	E	E
64	A fast animal ridden by jockeys	7-11yrs	horse	5531	2149	88%	8	7503	2.00	1	S
65	A building people live in	5-7yrs	house	5051	1636	79%	12	49424	-	2	S
66	A short coat	5-7yrs	jacket	7985	1645	53%	29	2800	2.00	4.00	S
67	A container for pouring liquids	7-11yrs	jug	6286	3062	65%	15	517	3.00	4	4
68	A jumping Australian animal with a pouch	7-11yrs	kangaroo	4610	2012	100%	0	114	4.00	S	2
69	A kitchen appliance for boiling water	7-11yrs	kettle	5415	1829	79%	12	847	3.00	3.00	3 and 4
70	A man with a crown	5-7yrs	king	5641	1423	95%	4	15793	3.00	1	1
71	An Australian bear with grey fur	7-11yrs	koala	4771	2543	69%	15	17	-	E	E
72	A metal or wooden object with rungs for climbing	11-16yrs	ladder	4608	2734	50%	23	1271	3.00	4	4
73	A baby sheep	5-7yrs	lamb	5073	1360	95%	4	1635	2.00	1	1
74	Green flat things that grow on trees and plants	11-16yrs	leaves	5057	3543	79%	12	6354	2.00	3	2 and 3
75	An African animal that roars	5-7yrs	lion	5833	2007	90%	9	1226	3.00	2	S
76	A small scaly reptile with four legs and a long tail	7-11yrs	lizard	3691	2984	69%	21	210	4.00	4	4
77	A large piece of wood for putting on the fire	7-11yrs	log	4320	2651	50%	23	1231	3.00	4	S
78	A grown-up boy	7-11yrs	man	5604	1668	60%	14	58821	2.00	3	S
79	A mythical woman with a fish tail	7-11yrs	mermaid	3620	1839	88%	8	89	3.00	3	3

80	A cheeky animal that likes swinging in trees	5-7yrs	monkey	6170	2130	100%	0	556	3.00	1	1
81	A rodent that likes cheese	7-11yrs	mouse	5669	1716	84%	13	1854	2.00	S	3
82	A large straight-sided cup without a saucer	11-16yrs	mug	5845	3254	29%	26	705	2.00	E	E
83	A bird's home which is made of twigs	5-7yrs	nest	5682	1453	95%	4	1426	3.00	1	1
84	A religious woman living in a convent	11-16yrs	nun	4415	1746	63%	10	418	-	E	E
85	A person who looks after ill or injured people	7-11yrs	doctor	3396	2323	65%	2	10224	-	E	E
86	A vegetable that makes you cry when you cut it	5-7yrs	onion	5246	2222	100%	0	675	3.00	S	1
87	A wise bird that flies at night	5-7yrs	owl	5335	2059	95%	4	176	3.00	2	S
88	A large black and white bear from China	7-11yrs	panda	7429	2337	94%	4	387	3.00	E	E
89	A brightly coloured bird that can talk	5-7yrs	parrot	4229	2307	89%	9	224	3.00	3	S
90	A bell-shaped pale green fruit	11-16yrs	pear	4744	3809	56%	19	507	3.00	4	4
91	A bird which lives in the South Pole	7-11yrs	penguin	6335	3846	72%	16	1916	3.00	S	3
92	A musical instrument with black and white keys	7-11yrs	piano	4076	1931	80%	7	1320	2.00	2	2
93	A pink farm animal with a snout	5-7yrs	pig	3986	1861	100%	0	3220	-	1	1
94	A person who flies a plane	5-7yrs	pilot	3822	1674	85%	13	270	3.00	3.00	3.00
95	A sailor with an eye-patch and parrot	5-7yrs	pirate	7010	2036	95%	4	480	3.00	2	2
96	Flat round dough with a savoury topping	11-16yrs	pizza	5630	2954	75%	11	3391	3.00	E	E
97	A flying vehicle with wings	7-11yrs	plane	5149	2157	56%	11	3965	2.00	S	3.00
98	A flat dish that you put food on	5-7yrs	plate	6083	2025	95%	4	865	2.00	2	2
99	A vegetable used to make chips	7-11yrs	potato	5566	2110	83%	8	73	4.00	3	3
100	A big orange vegetable for Halloween	7-11yrs	pumpkin	8944	2232	89%	9	474	2.00	3	3
101	A baby dog	5-	puppy	6221	1325	95%	4	1403	2.00	1	S

		7yrs									
102	A hopping animal with long ears	7-11yrs	rabbit	6319	3620	68%	7	2869	3.00	S	3
103	A large stone	7-11yrs	rock	6656	1675	58%	20	1026	3.00	4.00	4.00
104	A large cloth bag often used for coal or potatoes	7-11yrs	sack	5162	3186	73%	11	323	-	E	E
105	A metal cooking pot with a long handle	7-11yrs	pan	6247	2483	56%	15	1880	-	4	4
106	A huge water area covering most of the earth's surface	5-7yrs	sea	5228	2413	68%	16	1383	4.00	4.00	4.00
107	A furry sea animal with blubber and flippers	11-16yrs	seal	6590	3351	44%	23	2983	2.00	E	E
108	An animal that gives wool	5-7yrs	sheep	6631	1520	100%	0	1381	3.00	1	S
109	A piece of furniture to store books in	7-11yrs	bookshelf	5450	2208	47%	14	70	-	E	E
110	The hard case around a snail or tortoise	5-7yrs	shell	5791	2683	89%	9	2182	3.00	2	S
111	A piece of footwear with laces	5-7yrs	shoes	5728	3215	50%	10	1126	2.00	S	4
112	A reptile with venomous fangs	11-16yrs	snake	4818	2419	53%	23	730	3.00	4	S
113	A figure made in winter with a carrot nose	5-7yrs	snowman	4415	2017	100%	0	76	2.00	2	S
114	A long soft seat for many people	7-11yrs	sofa	5817	2531	70%	11	921	2.00	3	4
115	A man who works for the military	11-16yrs	soldier	4973	2425	88%	3	1725	3.00	E	E
116	A high seat without a back	7-11yrs	stool	4802	5110	93%	4	815	3.00	2	2
117	A kitchen appliance used for baking	7-11yrs	oven	8415	2981	53%	23	1311	3.00	4.00	4.00
118	A large bag for taking clothes on holiday	7-11yrs	suitcase	6820	2195	95%	4	580	3.00	2	2
119	A long-necked white bird often seen on lakes	5-7yrs	swan	5136	2981	59%	19	522	3.00	4	4
120	A water-filled structure to dive into	7-11yrs	swimming pool	5201	2016	56%	15	982	3.00	4	S
121	A seat hung on ropes that moves back and forth	7-11yrs	swing	4005	2859	86%	8	799	3.00	3	2 and 3
122	A piece of furniture with legs	5-7yrs	table	5690	2460	47%	19	1896	2.00	E	E

	and a flat top										
123	Something you turn to let water out	5-7yrs	tap	5751	2761	84%	8	2042	3.00	3	2
124	What you pour tea from	11-16yrs	teapot	5583	1791	50%	8	241	3.00	4	4
125	An outdoor camping shelter	5-7yrs	tent	7875	1753	81%	12	1097	3.00	3	3 and 4
126	A wild cat with orange and black stripes	5-7yrs	tiger	6731	2656	84%	8	888	3.00	S	3
127	An object in the bathroom that flushes	5-7yrs	toilet	7326	2456	95%	4	1540	2.00	1	1
128	A soft round red fruit eaten as a vegetable	7-11yrs	tomato	7516	2277	65%	20	720	3.00	4	4
129	Public transport on the railway	7-11yrs	train	4198	2464	88%	8	46	-	3	S
130	Canvas joined to a frame with springs for jumping on	7-11yrs	trampoline	3273	3656	83%	7	6136	2.00	E	E
131	A large plant with a trunk, branches and leaves	5-7yrs	tree	5867	2457	84%	3	1110	2.00	3	2 and 3
132	Another name for a lorry	7-11yrs	truck	6784	2020	88%	8	255	3.00	3.00	3.00
133	A slow creature with a shell and long life-span	7-11yrs	tortoise/snail	6489	3074	25%	10	7861	2.00	E	E
134	A mythical horse-like creature with a horn	7-11yrs	unicorn	5252	2926	72%	11	94	4.00	4	4
135	A partition made of bricks	11-16yrs	house	5787	2021	50%	18	11298	2.00	4	S
136	A cascade of water over the edge of rocks	11-16yrs	waterfall	5868	1833	57%	14	387	-	4	S
137	A huge sea mammal	11-16yrs	whale	6879	2100	81%	7	571	3.00	3	S
138	A see-through cup with a stem	11-16yrs	wineglass	5228	2686	44%	13	13	2.00	E	E
139	A woman who does magic and flies on a broomstick	5-7yrs	witch	5908	2710	100%	0	573	3.00	1	1
140	A horse-like animal with black & white stripes	5-7yrs	zebra	4613	2180	94%	4	222	3.00	S	2

3.3 Picture Describing

Item no.	**Dominant response** (modal subject-verb-object)	Child RT (ms)	Adult RT (ms)	Child RD (ms)	Adult RD (ms)	RA	H	Freq	Errors	AoA (likert)	LEVEL
1	ant-eating-cherry	1706	1432	5959	4448	68%	15.95	647	5%	3.00	3 and 4
2	baby-drinking-cup	1432	1442	9197	4428	50%	19.53	10285	10%	2.00	2 and 3
3	badger-eating-onion	1836	1846	6717	5425	84%	12.99	515	0%	3.00	2 and 3
4	bear-eating-pumpkin	1363	1233	5963	4229	95%	4.40	2835	0%	*3.50*	1
5	bird-drinking-fountain	1438	1389	5957	5208	61%	19.98	2272	10%	3.00	2 and 3
6	book-falling-shelf	1383	1250	5962	4641	86%	7.84	13607	30%	2.50	1
7	bowl-falling-fridge	1599	1293	6177	4650	93%	4.01	1629	25%	2.50	1
8	boy-jumping-bed	2003	1193	6283	4422	90%	8.80	13697	0%	1.50	1
9	bread-falling-basket	1906	1798	7519	4442	69%	15.33	2446	35%	2.50	4
10	camel-drinking-bath	1523	1143	7026	4242	75%	11.80	2074	0%	2.50	1
11	camera-falling-chair	1276	1255	6356	4420	100%	0.00	5044	20%	2.50	1
12	cat-drinking-jug	1142	1143	8394	4642	74%	11.60	2185	0%	2.00	1
13	clown-jumping-trampoline	1535	1080	6874	4438	100%	0.00	204	0%	2.00	1
14	cow-eating-hedge	1387	1405	5959	4450	55%	10.24	1148	0%	2.50	2
15	cowboy-drinking-glass	2502	1461	6242	4676	54%	18.40	4834	35%	2.50	4
16	crab-falling-hand	1629	1608	5979	4977	85%	2.94	16880	30%	2.50	1
17	kangaroo-jumping-fence	1498	1594	6179	4436	56%	6.25	891	5%	3.50	2
18	dolphin-jumping-hoop	1421	1424	5964	4430	90%	3.47	450	0%	3.00	1
19	donkey-eating-carrot	1372	1106	5977	4436	90%	3.47	444	0%	2.50	1
20	duck-eating-cake	1648	1394	5970	4434	90%	8.80	2036	0%	2.00	1
21	elephant-drinking-waterfall	1174	1343	5977	4669	88%	8.36	644	15%	2.00	3 and 4
22	fox-jumping-bus	1657	1872	5952	4850	80%	12.29	3804	0%	2.50	2
23	frog-jumping-stool	1506	1273	5983	4649	60%	2.06	654	0%	2.50	1 and 2

24	giraffe-eating-leaves-from-tree	1406	1278	5972	4459	100%	0.00	10353	55%	2.00	4
25	girl-jumping-swing	1682	1690	7060	4439	95%	4.33	5622	5%	2.00	1
26	gloves-falling-bag	1528	1325	6391	4441	85%	7.64	3023	35%	2.50	4
27	goat-eating-hat	1584	1333	5958	5588	76%	11.65	1756	10%	2.50	2 and 3
28	guitar-falling-sofa	1395	1226	5971	4432	88%	8.19	1839	20%	2.50	2 and 3
29	horse-jumping-gate	1244	1038	7067	4431	74%	11.18	5465	5%	2.50	2
30	jacket-falling-suitcase	1576	1789	5953	9618	47%	18.05	1661	25%	2.50	3 and 4
31	kangaroo-jumping-car	1404	1423	5961	4645	100%	0.00	13500	25%	3.00	1
32	teapot-falling-oven	1185	1114	5949	4657	22%	45.04	776	10%	3.00	4
33	key-falling-keyhole/lock	1489	1556	5960	4637	36%	4.78	4990	30%	-	3 and 4
34	king-drinking-cup	1343	1544	5985	4449	75%	6.58	13879	40%	2.50	4
35	koala-eating-pear	1173	1247	5959	4239	81%	12.30	121	5%	3.00	3 and 4
36	ladder-falling-house	1462	1171	5960	5172	100%	0.00	25348	45%	3.00	4
37	lamb-drinking-bottle	1723	1516	5971	4442	69%	15.54	2783	20%	1.50	3
38	lion-jumping-train	1402	1388	6953	13606	84%	7.74	4544	5%	2.50	1 and 2
39	lizard-eating-biscuit	1650	1742	6804	4427	70%	11.48	383	0%	3.00	2 and 3
40	man-falling-boat	1354	1452	5954	4224	82%	2.78	32044	15%	1.50	2 and 3
41	mermaid-drinking-shell	1424	1322	5959	13503	80%	7.14	1136	20%	3.00	2 and 3
42	monkey-eating-banana	1575	1286	5960	4683	90%	8.80	540	0%	2.50	2
43	mouse-eating-cheese	1418	1163	5954	4452	95%	4.40	2185	0%	2.00	1
44	nest-falling-tree	1412	1494	7010	4437	94%	4.09	3781	20%	2.50	2
45	nun-falling-plane	1538	1242	6413	4661	25%	30.42	1905	15%	3.00	4
46	doctor-drinking-glass	1416	1408	9347	4751	31%	35.00	5243	30%	2.50	4
47	owl-eating-tomato	1281	1279	5947	4448	100%	0.00	448	0%	3.00	1
48	panda-eating-bamboo	1608	1177	5958	4423	76%	16.74	337	10%	3.00	3 and 4
49	parrot-drinking-tap	1617	1072	5945	4439	92%	3.82	1215	35%	3.00	4
50	penguin-eating-fish	1631	1373	5963	4441	100%	0.00	5629	0%	2.50	1
51	piano-falling-truck	1346	1260	5965	4442	53%	14.53	1086	25%	3.00	3

52	pig-eating-pizza	1317	1082	7066	4440	95%	4.40	900	0%	2.50	1
53	policeman-drinking-coconut	1747	1658	6823	4435	50%	14.36	1175	0%	-	2 and 3
54	pirate-drinking-barrel	1570	1150	5952	4427	93%	4.01	565	20%	3.50	2 and 3
55	potato-falling-sack	1361	1084	7167	5445	47%	18.63	946	25%	2.50	3 and 4
56	puppy / dog-eating-shoe	1481	1799	6555	4453	42%	15.24	3157	5%	2.00	3 and 4
57	rabbit-jumping-log	1415	1716	7176	4444	83%	7.60	1317	5%	2.50	2
58	seal-jumping-rocks	1228	1353	6536	4435	73%	16.08	4027	10%	3.50	3 and 4
59	sheep-drinking-pan	1344	1378	5970	4458	39%	13.72	1653	10%	2.00	3 and 4
60	snake-eating-apple	1221	1134	5947	4644	100%	0.00	1670	0%	2.50	1
61	army man-drinking-wineglass	1912	1380	5950	4457	29%	16.26	13	20%	2.00	4
62	squirrel-jumping-tree	1513	1590	5970	4678	74%	11.18	3182	5%	2.00	2 and 3
63	swan-eating-grapes	1435	1416	5978	4448	85%	7.88	750	0%	3.00	1 and 2
64	teapot-falling-off-table	1543	1283	6688	4423	67%	10.72	9765	25%	2.50	2 and 3
65	tiger-jumping-tent	1734	1413	5961	4655	94%	4.25	993	0%	3.00	1 and 2
66	tortoise-drinking-plate	1497	1427	5956	4451	40%	14.36	1458	25%	2.50	4
67	unicorn-jumping-snowman	1264	1167	8165	4446	90%	8.80	85	0%	3.00	2
68	whale-jumping-sea	1520	1317	13995	4459	44%	9.87	6581	5%	3.00	4
69	witch-jumping-fire	1587	1023	5964	5960	95%	4.40	7074	0%	2.50	1
70	zebra-drinking-pool	1630	1121	6138	4667	45%	15.44	2350	0%	2.50	4

4. THE NATIONAL STATISTICS SOCIO-ECONOMIC CLASSIFICATION (NS-SEC)

Please select your answer by ticking the appropriate box or, if completing a computerised version of this form, by underlining the appropriate answer.

Education history

At what age did you finish your full-time education?
Finished education at age of _____years
Not yet finished
Never went to school

Final qualification gained: _____

Employment status
The following questions refer to your current main job, or (if you are not working now) to your last main job.
Please select only one answer per question.

Employee or self-employed

Do (did) you work as an employee or are (were) you self-employed?

Employee ☐

Self-employed with employees ☐

Self-employed / freelance without employees (go to question 4) ☐

Number of employees (Employees)

For employees: indicate below how many people work (worked) for your employer at the place where you work (worked).
For self-employed: indicate below how many people you employ (employed). Go to section E when you have completed this question.

1 to 24 ☐

25 or more ☐

Supervisory Status

Do (did) you supervise any other employees?
A supervisor or foreman is responsible for overseeing the work of other employees on a day-to-day basis

Yes ☐

No ☐

Occupation

Please select one answer to show which best describes the sort of work you do. (If you are not working now, please select the answer to show what you did in your last job). SELECT ONE ANSWER ONLY

Modern professional occupations
such as: teacher - nurse - physiotherapist - social worker - welfare officer - artist - musician - police officer (sergeant or above) - software designer ☐

Clerical and intermediate occupations
such as: secretary - personal assistant - clerical worker - office clerk - call centre agent - nursing auxiliary - nursery nurse ☐

Senior managers or administrators ☐

(usually responsible for planning, organising and co-ordinating work and for finance)
such as: finance manager - chief executive 3

Technical and craft occupations
such as: motor mechanic - fitter - inspector - plumber - printer - tool maker - electrician - gardener - train driver ☐ 4

Semi-routine manual and service occupations
such as: postal worker - machine operative - security guard - caretaker - farm worker - catering assistant - receptionist - sales assistant ☐ 5

Routine manual and service occupations
such as: HGV driver - van driver - cleaner - porter - packer - sewing machinist - messenger - labourer - waiter / waitress - bar staff ☐ 6

Middle or junior managers
such as: office manager - retail manager - bank manager - restaurant manager - warehouse manager - publican ☐ 7

Traditional professional occupations
such as: accountant - solicitor - medical practitioner - scientist - civil / mechanical engineer ☐ 8

5. FAMILY HANDEDNESS QUESTIONNAIRE

Please complete this short questionnaire along with your child, and return it to Louise Croft (louise.croft.09@ucl.ac.uk) as soon as possible. To answer the questions please delete as appropriate, leaving only one answer.

A. Your child's hand preference

1. Do you consider your child to be left-handed or right-handed? **LEFT HANDED / RIGHT HANDED**

2. Do you consider your child to be strongly or moderately left/right-handed? **STRONGLY / MODERATELY**

If moderately:

3. Are there any tasks your child is able to do equally well with their right hand and with their left hand?

YES / NO

Which tasks? _____

4. Did your child ever suffer an injury which made it difficult or even impossible to use either hand for a long period of time? **YES / NO**

If so, which hand? **LEFT / RIGHT** For how long? _____

B. Hand preference of Family members:

Mother: **RIGHT / LEFT / AMBIDEXTROUS**

Father: **RIGHT / LEFT / AMBIDEXTROUS**

Siblings: **RIGHT / LEFT / AMBIDEXTROUS**

Grandparents (mother's side): **RIGHT / LEFT / AMBIDEXTROUS**

Grandparents (father's side): **RIGHT / LEFT / AMBIDEXTROUS**

6. IN-SCANNER EXPERIENCE QUESTIONNAIRE

Post-scan questionnaire for participants

Please read each sentence and then circle one answer from the options below. Please circle only one answer each time. There are no right or wrong answers.

A. The next 6 questions are asking you about how you felt while you were having your brain scan. (Try to remember exactly how you felt when you were lying inside the real brain scanner, playing the Panda Games).

1. **Did you feel calm inside the brain scanner?**

No | A little | Yes | Yes (very much)

2. **Did you feel nervous inside the brain scanner?**

No | A little | Yes | Yes (very much)

3. **Did you feel upset inside the brain scanner?**

No | A little | Yes | Yes (very much)

4. **Did you feel relaxed inside the brain scanner?**

No | A little | Yes | Yes (very much)

5. Did you feel <u>happy</u> inside the brain scanner?

No  A little  Yes  Yes (very much) 

6. Did you feel <u>worried</u> inside the brain
7. scanner?

No  A little  Yes  Yes (very much) 

If there is anything else you would like to say about <u>how you felt inside the brain scanner</u> you can write about it here:

_____

_____

_____

B. The next 4 questions are asking you about how <u>comfortable</u> it was inside the brain scanner. Please circle one answer each time.

1. Inside the scanner I felt…

Very comfortable  Comfortable  Uncomfortable  Very uncomfortable 

2. Did you mind being in a small space?

No (not at all) | A little | Yes | Yes (very much)




3. How long do you think you were inside the brain scanner?

A very long time (too long) | A long time (but it was ok) | Not very long | A short time




4. The headphones I wore were…

Very comfortable | Comfortable | Uncomfortable | Very uncomfortable




If there is anything else you would like to say about <u>how comfortable you felt</u> inside the brain scanner you can write about it here:

_____

_____

_____

C. The next 7 questions are asking <u>how much you enjoyed playing the Panda Games</u> (inside the scanner and at home). Please circle one answer each time.

1. Did you enjoy playing the Panda Games?

No (not at all) A little Yes Yes (very much)




2. **Did you get bored while you were playing the Panda Games inside the scanner?**

No (not at all) A little Yes Yes (very much)




3. **My <u>favourite</u> Panda Game was** (circle one game)…

Piicture naming

Picture describing

Listen and guess

Read and guess

Colour naming (words)

Colour naming (sentences)

Alien game

Face finder

4. **My <u>least favourite</u> game was** (circle one game)…

Picture naming

Picture describing

Listen and guess

Read and guess

Colour naming (words)

Colour naming (sentences)

Alien game

Face finder

5. How <u>difficult</u> did you find the Panda Games?

Very easy | Quite easy | Difficult | Very difficult

6. Did you find it <u>helpful to practise</u> the Panda Games at home, before your brain scan?

No (not at all) | A little | Yes | Yes (very much)

7. Did you find the <u>instruction booklet helpful</u>?

No (not at all) | A little | Yes | Yes (very much)

If there is anything else you would like to tell us about the Panda Games you can write it here:

D. **The next 4 questions are asking you about <u>your experience with the researcher</u> on the day of your brain scan.** Please circle one answer each time.

1. **How <u>clear</u> was the researcher about <u>what was going to happen</u> throughout the whole day…**

Not clear at all **Quite clear** **Clear** **Very clear**




2. **Did the researcher make you feel <u>safe</u>?**

No **A little** **Yes** **Yes (very much)**




3. **Was the researcher <u>friendly</u>?**

No (not at all) **A little** **Yes** **Yes (very much)**




4. **Would you have preferred it if you could have met the researcher <u>before</u> coming to the centre for testing?**

No **A little** **Yes** **Yes (very much)**




If there is anything else you would like to say about the researchers you met, you can write it here:

_____

_____

_____

_____

E. The next 3 questions are asking you about <u>your experience taking part in a scientific study</u>. Please circle one answer each time.

1. **What was the main reason you decided to take part in this study?** (please circle one)

For a picture of my brain

For the £10 reward

To help other people

To help with scientific research

Other

If you answered '**Other**' please tell us about the reason you decided to take part:

_____

_____

2. **What would you <u>change</u> about your day visiting us?**

_____

_____

_____

3. **Would you like to participate in a brain scan study again?** Please circle your answer.

Yes No

F. What is your ethnic origin? Please tick as appropriate.

☐ **White**

- ☐ British
- ☐ Irish
- ☐ Any other white background (please specify): ……

☐ **Mixed**

- ☐ White and Black Caribbean
- ☐ White and Black African
- ☐ White and Asian
- ☐ Any other mixed background (please specify): ……

☐ **Asian or Asian British**

- ☐ Indian
- ☐ Pakistani
- ☐ Bangladeshi
- ☐ Any other Asian background (please specify): ……

☐ **Black or Black British**

- Caribbean
- ☐ African
- ☐ Any other Black background (please specify): ……

☐ **Chinese**

☐ **Any other ethnic group** (please specify): ……

7. FACTORS UNDERLYING IN-SCANNER EXPERIENCE

In order to identify and validate constructs within the *In-scanner experience* questionnaire, I performed principal component analysis (PCA) using a varimax rotation to identify underlying orthogonal factors which explained variance in participant responses. I performed PCA with 25 iterations and extracted all variables with an eigenvalue >1. The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.59, confirming the sample size was sufficient for reliable factor extraction. PCA extracted three components explaining 34%, 29% and 12% of the variance respectively, and 75.27% of the variance in total (Table 31). I retained all variables with factor loadings >0.72, as per the recommendation by Stevens (1992) for samples with N<50.

	Component		
	1	2	3
Factor name and items	**In-scanner experience**	**Researcher interaction**	**Headphone comfort**
Task enjoyment: How difficult did you find the Panda Games?	0.796		
Task enjoyment: Did you enjoy the Panda Games?	0.736		
Physical comfort: Did you mind being in a small space?	0.784		
Physical comfort: How comfortable did you feel inside the scanner?	0.783		
Physical comfort: How long do you think you were inside the brain scanner?	0.724		
Researcher interaction: How clear was the researcher about what was going to happen on the day?		0.891	
Researcher interaction: Did the researcher make you feel safe?		0.878	
Researcher interaction: Was the researcher friendly?		0.871	
Physical comfort: Were the headphones comfortable?			0.96
Variance explained	**34.17%**	**29.38%**	**11.72%**

Table 31: Factor loadings on three components extracted with Principal Component Analysis, accounting for 75.27% of the variance in post-scan questionnaire responses in 24 children.

Principal component analysis did not identify the three factors I hypothesised; 1) Physical comfort, 2) Researcher interaction and 3) Task enjoyment. Instead, component 1 combined three *Physical comfort* and two *Task enjoyment* items. I therefore interpreted component 1 as an *In-scanner experience* factor. Component 2 was formed of all three items relating to *Researcher interaction*. Component 3 included one *Physical comfort* item, and from here on is referred to as the

Headphone comfort factor. To investigate the influence of these three variables on measures of scan success, I produced composite scores for each factor by calculating the mean rating across items on each factor (Table 32).

Measure	Mean	SD	Min.	Max.	Higher scores indicate…
In-scanner anxiety	28.19	7.68	20	43	More anxiety
In-scanner experience composite	3.18	0.57	1	4	Better experience inside the scanner
Researcher interaction composite	3.83	0.33	3	4	More positive interaction with the researcher
Headphone comfort	3.21	0.93	1	4	Headphones were more comfortable

Table 32: Four measures of in-scanner experience obtained from a post-scan questionnaire in a subgroup of 24 healthy children and adolescents.

8. ‘HOW TO PLAY’: THE PANDA GAMES ONLINE INSTRUCTION BOOKLET

THE PANDA GAMES

HOW TO PLAY

5-7 YEARS

CONTENTS

This booklet explains what you can expect when you have your brain scan. It also explains how to play *THE PANDA GAMES*. Please read this booklet before your language brain scan appointment. Please also make sure you have a go at the practice version of the Panda Games on your computer at home.

Super-Panda's Brain Scan Adventure

This is a story about Super-Panda who, like you, is going to get a brain scan!

One day, Super-Panda received a special package in the post. A letter inside the package invited Super-Panda to go and get **a special photo taken of his brain, called a brain scan**.

During his brain scan Super-Panda could play the Panda Games. The Panda Games were some fun speaking and listening games.

Super-Panda and his Mum had a go at all of the Panda Games on their computer at home.

They were really fun!

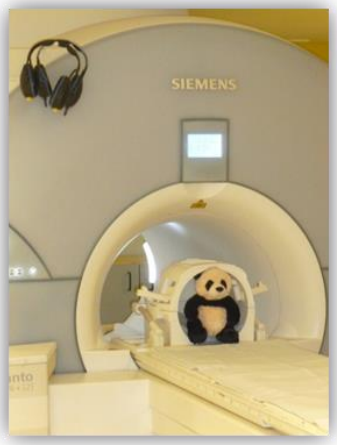
Super-Panda and his Mum got the train to **Great Ormond Street Hospital** in London. When they arrived they met a scientist. "First of all we are going to practise playing the Panda Games inside a **pretend brain scanner**!" she said.

Super-Panda inside the pretend brain scanner

Super-Panda climbed inside the pretend brain scanner. He was wearing some headphones so he could listen to the noises the real brain scanner would make. The sounds went chug-chug-chug and buzz-buzz-buzz. The scientist explained that it was really important for Super-Panda to lie as still as a statue inside the brain scanner and to speak very gently, in a whisper. **"If you move inside the brain scanner, the photo of your brain will be all fuzzy!" she said.**

Super-Panda's brain scan adventure

Once Super-Panda had practised in the pretend brain scanner, they went over to the room with the real brain scanner. Super-Panda and his Mum met the radiographer. **A radiographer is a person specially trained to take photos using a brain scanner**. Together they filled in some forms and then went in to see the real brain scanner.



The radiographer helped Super-Panda onto the little table which stuck out of the brain scanner. She put some headphones on Super-Panda and made sure he was comfortable. "Ready?" she asked. "Ready!" said Super-Panda. The radiographer pressed a button on the side of the brain scanner and the table moved inside the brain scanner. **"This is like going into a spaceship!" said Super-Panda.**

Once he was inside the brain scanner, the radiographer went into a different room and put on a DVD, which played on the screen inside the brain scanner. It was one of Super-Panda's favourite films, he had brought it with him! The radiographer then began to take some photos of Super-Panda's brain whilst he watched his DVD, and **the scanner began to make those funny noises he had heard earlier…chug-chug-chug, buzz-buzz-buzz**.

Super-Panda inside the real brain scanner

Super-Panda's Brain Scan Adventure

Once those photos were finished, the radiographer said "Now then Super-Panda, are you ready to play the Panda Games?". "Yes!" said Super-Panda. The radiographer told Super-Panda which of the Panda Games he was going to play, one at a time. **Super-Panda tried as hard as he could when he played the Panda Games**, **so that he got the best photo of his brain!** To make sure his photo was nice and clear, **he lay as still as possible and whispered his answers gently**, just like he had practised!

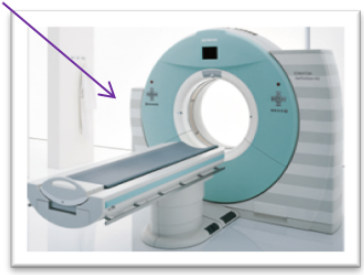
Once Super-Panda had played all of the Panda Games, and after the radiographer had helped him out of the scanner, the scientist said "Well done Super-Panda, you did so well! I bet the photo of your brain is going to be great. Did you know, you will be able to see all the bits of your brain that were busy working when you played the Panda Games. They light up like a Christmas tree!".

Super-Panda and his Mum said goodbye and went home on the train. A few days later Super-Panda's Mum received an email from the scientist. "The photo of your brain has arrived!" she said. "Wow, look at all the bits that lit up!" Super-Panda said. He was so excited he printed the photo of his brain and took it into school the next day.

Everyone was very excited about Super-Panda's brain scan !

Your brain scan

You are going to have a language brain scan! This means you will lie inside a **brain scanner**



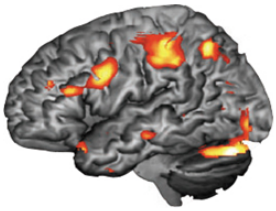
The brain scanner looks like a doughnut. You lie in the middle bit. It is **open at both ends**.

There is a TV screen inside the scanner, which you will be able to see. You might even be able to watch a DVD for part of the time you are in the scanner! Why don't you bring your favourite DVD with you?

The brain scanner is a **big camera**, which takes photos of your brain!

While you are lying inside you get to play *The Panda Games!*

The Panda Games are fun to play. They help to show us what your brain looks like when you talk or listen.



When you play the Panda Games inside the scanner, **the parts of your brain that are working light up**, **and we take lots of photos of your brain in action!**

You can watch a video to see what it is like to have a brain scan at Great Ormond Street Hospital by clicking here.

Keeping still inside the scanner

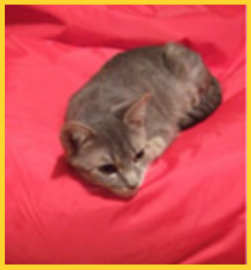
When you are lying inside the brain scanner it is very important that you keep <u>as still as possible</u>!

This is because we are taking lots of photos of your brain.

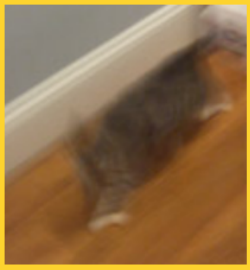
<u>If you move, the pictures will be fuzzy!</u>

This means we will have to do it again.

Look at these examples… →



This cat is very still and the photo is very clear.



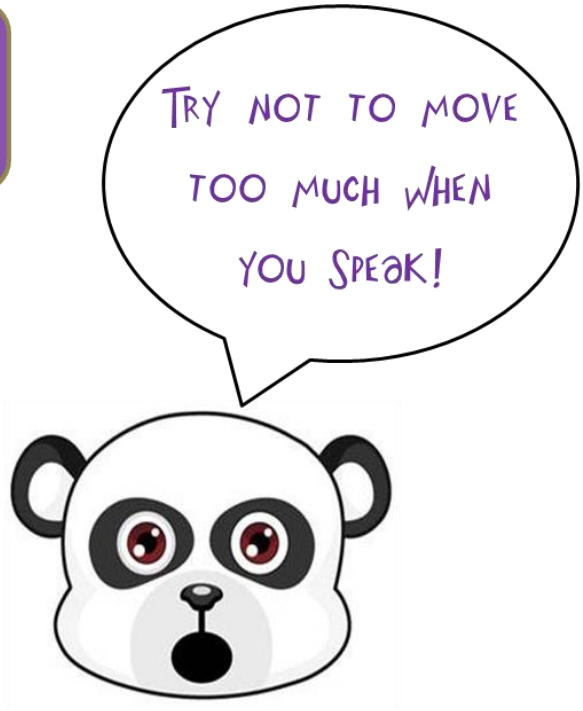
This cat is moving and the photo is very fuzzy.

Speaking inside the scanner

When you play *The Panda Games* inside the brain scanner, you have to say your answers out loud.

Try to **speak softly** inside the brain scanner, like a **whisper**.

This is important to help you **keep your head still**, and to make sure the photo is really clear.



Before your brain scan, you can practice whispering answers out loud in our **pretend brain scanner!**

The Panda Games

There are 6 games you may play…

- Picture naming
- Picture describing
- Listen and guess
- Colour naming (words)
- Colour naming (sentences)
- Alien game

Game instructions

This booklet explains how the games work, with **examples**, **game rules** and **answers to some questions** you may have. Please make sure you read this before your appointment. Perhaps your parents can go through this with you?

Practicing at home

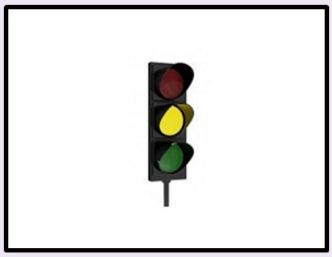
You can play each game online, at home, by clicking the yellow and purple boxes above. Having a go at home will make it easier for you to do well when you play the games inside the brain scanner! To play the games online you need to enter the following information:

- Password: **super-panda**
- Email: **panda@games.com**
- Name: Enter your **participant ID**

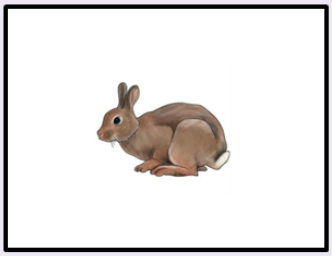
Make sure the sound on your computer is turned on, and the volume is up!

How the games work...

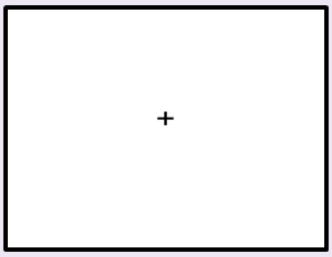
At the beginning of each game you will see a yellow traffic light.
This tells you to get ready!



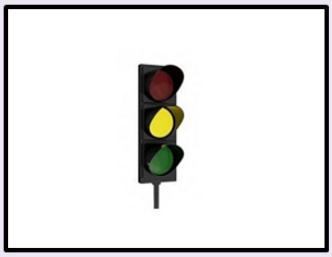
Then the pictures or sounds will begin to appear, one at a time, and you can say your answer.



In-between each go, you will see a little cross. You can relax while this is on the screen. Sometimes the cross will stay on the screen for quite a long time. This is ok, just relax! This is a nice break for you!



Before the pictures or sounds begin again, you will see the traffic light.
This tells you to get ready!



Then the pictures or sounds will start again, and so on!

Each game takes less than 5 minutes to play!

Picture Naming

You will see a picture on the screen inside the brain scanner. You have to try to name the picture - say the first thing that pops into your head!

For example, if you saw this picture...

...you might say 'Car'

What is this for?

This game helps us to see which parts of your brain are involved in saying names.

Game rules...

- Make sure you say the **name** only. Nothing else. So try not to say 'ummmm car' or 'it's a car'. **Just say 'car'.**

What if I don't know what the picture is?

That's ok! Just have a guess, or you can say '**don't know**' .

What if my answer is different to someone else's?

That's ok, just say what **you** think it is! There are no wrong answers!

Picture describing

You will see a picture on the screen inside the brain scanner. You have to say what is happening in the picture, in one short sentence.

For example, if you saw this picture...

...you might say "The dragon is jumping from the helicopter"

What is this for?

This game helps us see which parts of your brain are used to say sentences.

Game rules...

- Try to say the **whole sentence** like in the example. Try <u>not</u> to say anything shorter like "The dragon is jumping" or just "Jumping".
- There are only **4** actions in the pictures - **eating, drinking, jumping or falling**. Try to say one of these action words if you can.
- Say it as if it is happening **now**. For example "The pen **is falling** from the table". Try <u>not</u> to say "The pen **fell** from the table".
- If you see something drinking, try to say **what it is drinking from**.

For example, here you might say **"The nurse is drinking <u>from a glass</u>"**

What if I can't tell what is happening in the picture?

That's ok! Just say what action you think fits best - eating, drinking, jumping or falling!

If you really cant guess, say 'I don't know'!

Listen + Guess

You will be wearing headphones inside the brain scanner. In these headphones you will hear a description of something and you have to try to guess what is being described!

For example, if you heard...

"Something colourful with petals"

...you might say 'Flower'

What is this for?

This game shows which parts of your brain help you listen carefully.

Game rules...

- Make sure you say the **NAME** only. Nothing else. So try **not** to say "ummmm flower" or "it's a flower". **Just say "flower".**

What if I don't know what the answer is?

Just try your best! If you can't think of any answer, just say '**don't know**'!

What if my answer is different to someone else's?

That's ok! Just say what **you** think it is! There are no wrong answers!

Read + Guess

This game is the same as the other guessing game, but instead of hearing descriptions you will have to **read** them. (If you find the practice version of this game difficult, tell the researcher and they can choose an easier version for you.)

For example, if you read…

"Something you use to write with"

…you might say 'pen'

What is this for?

This game helps us see which parts of your brain help you read.

Game rules…

- Make sure you say the **NAME** only. Nothing else. So try **not** to say "ummmm pen" or "it's a pen". **Just say "pen".**
- Make sure you **read the sentence in your head**, try not to read it out loud. Only say your answer out loud.
- Try to **keep your lips still** while you read the sentence in your head.
- **Try not to make any noises while you read**, like saying 'ummmm' or 'errrr'.

What if I don't know what the answer is?

That's ok! Just try your best! If you can't think of any answer say '**don't know**'!

What if my answer is different to someone else's?

That's ok too! Just say what **you think it is!**

What if I can't read a word?

That's ok! Try your best to read as much of the sentence as you can and make your best guess!

Colour Naming - Words

You will see a funny pattern on the screen. Around this pattern is a coloured square. It will be yellow, blue, green or red. You have to name the colour of the square.

For example, if you saw this picture...

...you might say "yellow"

What is this for?

This game allows us to find out which parts of your brain move your mouth and tongue when you say a word.

Game rules...

- Make sure you say **one** colour name only.
- Try not to say "ummmm yellow"". **Just say "yellow".**

What if I don't know what the colour is?

That's ok! Just say what colour you think it is most like.

What if my answer is different to someone else's?

That's ok too! Just say what **you** think it is!

Colour Naming - Sentences

This game is like the other colour naming game. You will see a funny pattern on the screen, with a coloured square around it. This time, you have to say **"The colour of the square is..."** and then say the colour.

For example, if you saw this picture

...you might say **"The colour of the square is blue"**

What is this for?

This game lets us see which parts of your brain control your mouth and tongue when you say whole sentences.

Game rules...

- Make sure you say **the whole sentence exactly the same each time** (apart from the colour of course!)

What if I don't know what the colour is?

That's ok! Just say what colour you think it is most like.

What if my answer is different to someone else's?

That's ok too! Just say what **you** think it is!

Alien Game

You will be wearing headphones inside the brain scanner. In these headphones you will **hear some aliens talking** (you won't be able to understand what they say because you can't talk alien language!). You have to try to **guess whether each alien is a boy or a girl**, based on how they sound. Make sure you listen to the examples (by clicking **here!)** - Can you tell the difference?

If you heard an alien that sounded like a girl...

...you would say 'girl'

If you heard an alien that sounded like a boy...

...you would say 'boy'

What is this for?

This games lets us see which parts of your brain help you hear.

Game rules...

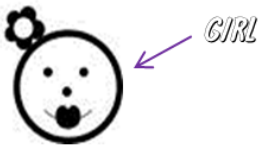
- Make sure you say only "boy" or "girl" each time. Nothing else. So try **not** to say "ummmm boy" or "it's a boy". Just simply say "boy".
- Although the aliens might sound pretty funny, **try your best not to laugh...**remember what we said about staying still!

What if I don't know what the answer is?

That's ok! Just make your best guess!

Face Finder

You will see a picture on the screen inside the brain scanner. The picture will show **some funny symbols**. Somewhere hidden in between the symbols is a **boy or a girl's face,** like this…



The boy is wearing a hat and has a bushy moustache!

The girl has a flower in her hair and is wearing lipstick!

You have to look along the symbols (as if you were reading) to find the face. Once you find the face, say if you think it is a boy or a girl!

For example, if you saw this...

...you might say "girl"

What is this for?

This shows us which parts of your brain are used for seeing.

Game rules:

- Make sure you say **"boy" or "girl" only**. Nothing else.
- Remember to **look along all the symbols** from left to right, like reading.

What if I can't tell whether it is a boy or a girl?

That's ok! Just try your best! If you can't see it very well, **tell the researcher!**

Word Game

You will be wearing headphones inside the brain scanner. In these headphones you will **hear a noun** (an object or animal name). You have to **say a doing word** (a verb) that goes with that object/animal. Every so often you will hear a funny muffled noise – you can just relax when you hear this.

You can have a go at this game online by clicking here.

For example, if you heard 'chair'

…you might say 'sit'

If you heard 'ball'…

…you might say 'kick'

What is this for?

This games lets us see which parts of your brain help you think of different words

Game rules…

- Make sure you say only 1 doing verb each time.
- Try **not** to say "ummmm" when you are thinking of an answer
- If you don't know the answer, just say 'don't know' or stay silent

What if I don't know what the answer is?

That's ok! Just make your best guess or say 'don't know'!

Questions and further information

We hope you enjoy practising *THE PANDA GAMES!*

If you have any questions please speak to Louise. She works in the research team and can answer any questions you might have. Her contact details are below.

Louise Croft

Tel: 0207 905 2729

E-mail: louise.croft.09@ucl.ac.uk

9. QUALITY RATING CRITERIA

Images A-G show individual participant scans during the Listen and Name Game (left is left, right is right), compared to rest. Results are shown at three statistical thresholds:

Red = $p<0.05$, FWE corrected

Green = $p<0.001$, $k>20$ (uncorrected)

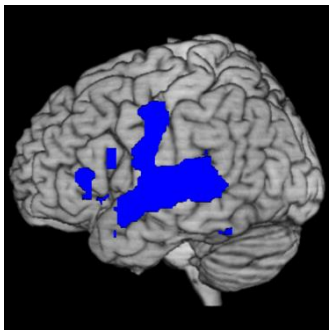
Yellow = $p<0.01$, $k>20$ (uncorrected)

Please rank images A-G from 1 to 7, where 1 is the highest quality image and 7 is the poorest quality image.

Please assess quality according to the following criteria:

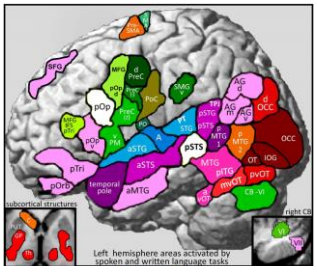
- How much of the target network (Figure 1) or its right hemisphere homologues, or biologically plausible regions (Figure 2) are activated and at what threshold.
- How much noise there is in the image:
    - Activation in biologically implausible regions (Figure 2), especially at a low threshold (yellow)

Figure 1: Activation of the auditory comprehension network (green) in the left



hemisphere.

Figure 2: The language system and it's functional subunits (from Price 2012). According to this functional-neuroanatomical model, the auditory comprehension network should involve pink, blue, green and white regions.



10. SAMPLE CHARACTERISTICS FOR HEALTHY CHILDREN INCLUDED IN GROUP MAP ANALYSES (CHAPTERS 6-7)

10.1 Picture naming

10.1.1 Age

Age groups were well-matched for demographics, neuropsychological performance, task practice and in-scanner movement (Table 16). There was a trend suggesting maximum in-scanner movement differed between age groups for CNw (Table 16). However, further analyses showed no significant correlation of maximum movement with age (p=0.99).

10.1.2 Sex

There were no significant differences in age, socio-economic status, neuropsychological performance or in-scanner movement between males and females. However, the Kruskal Wallis test for non-parametric data showed a significant difference in practice time for PN (X^2(1)=8.26, p=0.004) and CNw (X^2(1)=5.88, p=0.02), such that females practiced PN (mean=3.16mins, SD = 2.45) and CNw (mean=1.59mins, SD= 0.31) significantly more than males (mean=0.96mins, SD=1.61, and mean =0.31mins, SD=0.67, respectively).

10.1.3 Task (versus baseline)

Wilcoxon Signed-Ranks test showed participants practiced PN longer than CNw (Z=-3.57, p<0.001). There was no significant difference in movement between tasks (p values > 0.43).

10.2 Picture describing

10.2.1 Age

Age groups were well-matched for demographics, verbal ability, practice time, and in-scanner movement (Table 17). However, nonverbal ability (PIQ) differed significantly between age groups (Table 17). Post hoc independent samples t tests showed participants in early childhood had significantly higher

non-verbal abilities compared to participants in late childhood ($t(15)=2.21$, $p=0.04$; mean difference = 15.53 standard score points) and adolescence ($t(15)=3.26$, $p=0.005$; mean difference = 22 standard score points). Similarly, there was a trend for a difference in executive functioning (Table 17), such that younger participants had better general executive skills for their age relative to older children and adolescents.

10.2.2 Sex

There were no significant differences in sample characteristics between males and females (all p values >0.18).

10.2.3 Task (versus baseline)

Wilcoxon Signed-Ranks test showed significantly increased maximum movement during PD compared to CNs ($Z=-2.19$, $p=0.03$), suggesting participants moved more during PD than during its baseline task. However, mean movement for each task was minimal (well below 1 voxel). Participants practiced PD more than CNs ($Z=-2.60$, $p=0.009$).

10.3 Listen and Name

10.3.1 Age

Age groups were well matched for demographics, neuropsychological performance, practice time and mean in-scanner movement (Table 18). There was a trend suggesting maximum movement during AG differed between age groups (Table 18). However, there was no significant correlation of in-scanner movement with age for either task (p values >0.22).

10.3.2 Sex

There were no significant differences in demographics, neuropsychological performance or in-scanner movement between females and males (p values >0.32). There was a trend for a difference in practice time between males and females for LN ($p=0.07$) and AG ($p=0.09$) such that females practiced both tasks for more time compared to males.

10.3.3 Task (versus baseline)

There were no significant differences in task practice time or in-scanner movement between LN and AG (p values > 0.46)

10.4 Read and Name

10.4.1 Age

Age groups were well matched for demographics, neuropsychological performance, pre-scan practice time and in-scanner movement (Table 19). Although children had significantly higher core language ability compared to adolescents, the difference was minimal and not clinically significant (mean difference = 2 standard score points). Encouragingly, variables specifically predictive of reading proficiency (phonological awareness and rapid naming) were well-matched across age (Table 19).

10.4.2 Sex

There was no significant difference in practice time between males and females for either task, and no significant difference in movement for FF (p values > 0.39). There was a trend for increased maximum (p=0.09) and mean (p=0.08) movement during RN for females compared to males. However movement was minimal across the entire sample for both tasks (Table 19).

10.4.3 Task (versus baseline)

There was no significant difference in practice time between RN and FF (p=0.80). There was a trend suggesting higher mean (p=0.07) and maximum (p=0.08) movement during FF compared to RN. However, movement was minimal for both tasks (well below 1 voxel).

11. GROUP MAPS RESULTS: SIGNIFICANT ACTIVATION PEAKS

Montreal Neurological Institute (MNI) co-ordinates for significant activation peaks (>8mm apart) are listed for each contrast performed in Chapter 5, along with *t*, *z*, and *p* values, and notes on the approximate localisation of activation based on the Automatic Anatomic Labels (Tzourio-Mazoyer et al., 2002) provided by the xjview toolbox (http://www.alivelearn.net/xjview8/). Peaks are listed according to activation strength (*t* value), from strongest to weakest, at two statistical thresholds (labelled).

11.1 Picture Naming (PN)

Picture Naming (>rest)						
Peak values			MNI co-ordinates			Localisation notes
t	*z*	*p* (FWE)	x	y	z	
***p* <0.05 (FWE corrected)**						
17.46	65535.00	0.00	27	-91	-4	Right posterior inferior occipital gyrus
15.06	65535.00	0.00	36	-85	-4	Right posterior inferior occipital gyrus
11.61	7.39	0.00	42	-73	-10	Right posterior inferior occipital gyrus
12.83	7.75	0.00	-24	-94	-4	Left posterior inferior occipital gyrus
8.92	6.40	0.00	-36	-40	-22	Left vOT
7.88	5.94	0.00	-42	-79	-10	Left posterior inferior occipital gyrus
7.72	5.86	0.00	30	-1	-34	Right anterior fusiform/parahippocampal gyrus
7.48	5.74	0.00	24	-1	-22	Right amygdala
6.66	5.32	0.00	-54	-7	50	Left precentral gyrus
6.14	5.02	0.01	-57	-1	20	Left precentral gyrus (ventral)
5.88	4.87	0.01	-42	-13	38	Left premotor cortex - medial, within IFS
5.67	4.75	0.02	-51	-7	29	Left precentral gyrus - medial, within IFS
5.63	4.72	0.02	45	-10	35	Right precentral gyrus - medial, within IFS
5.38	4.56	0.04	-6	-34	-31	Left midbrain
***p* <0.001, *k* >10 (uncorrected)**						
17.46	65535.00	0.00	27	-91	-4	Right posterior inferior occipital gyrus
15.06	65535.00	0.00	36	-85	-4	Right posterior inferior occipital gyrus
12.83	7.75	0.00	-24	-94	-4	Left posterior inferior occipital gyrus
6.66	5.32	0.00	-54	-7	50	Left precentral gyrus
6.14	5.02	0.01	-57	-1	20	Left precentral gyrus (ventral)
5.88	4.87	0.01	-42	-13	38	Left premotor cortex - medial, within IFS
5.63	4.72	0.02	45	-10	35	Right precentral gyrus - medial, within IFS
4.53	3.99	0.29	54	-1	23	Right precentral gyrus
4.41	3.91	0.37	63	2	17	Right precentral gyrus
4.94	4.27	0.12	-3	5	62	Left SMA
4.33	3.85	0.43	-6	23	65	Left SMA

Colour Naming (>rest)						
Peak values			MNI co-ordinates			Localisation notes
t	*z*	*p* (FWE)	x	y	z	
***p* <0.05 (FWE corrected)**						
10.47	7.00	0.00	27	-94	-1	Right posterior inferior occipital gyrus
10.44	6.99	0.00	-24	-97	-1	Left posterior middle occipital gyrus
7.49	5.75	0.00	-33	-91	-13	Left posterior lingual gyrus
6.46	5.20	0.00	39	-64	-10	Right posterior inferior occipital gyrus
6.07	4.99	0.01	51	-7	29	Right precentral gyrus
5.87	4.87	0.01	-33	-85	-22	Left posterior inferior occipital gyrus
5.74	4.78	0.02	39	-55	-22	Right vOT
5.49	4.63	0.03	33	-73	-19	Right lingual gyrus
5.47	4.62	0.04	-54	-10	44	Left precentral gyrus
5.39	4.57	0.04	24	-10	-25	Right hippocampus (head)
***p* <0.001, *k* >10 (uncorrected)**						
10.47	7.00	0.00	27	-94	-1	Right posterior inferior occipital gyrus
10.44	6.99	0.00	-24	-97	-1	Left posterior middle occipital gyrus
7.49	5.75	0.00	-33	-91	-13	Left lingual gyrus
6.46	5.20	0.00	39	-64	-10	Right posterior inferior occipital gyrus
6.07	4.99	0.01	51	-7	29	Right precentral gyrus
5.74	4.78	0.02	39	-55	-22	Right vOT
5.47	4.62	0.04	-54	-10	44	Left precentral gyrus
5.39	4.57	0.04	24	-10	-25	Right hippocampus (head)
5.33	4.53	0.05	-51	-10	26	Left precentral gyrus
5.08	4.36	0.09	-48	-13	38	Left precentral gyrus
4.42	3.91	0.39	12	8	68	Right SMA
4.12	3.70	0.62	27	-10	-10	Right hippocampus (head)
3.89	3.52	0.81	-30	-10	-31	Left anterior fusiform gyrus

Semantic processing from visual input (PN>CNw)						
Peak values			MNI co-ordinates			Localisation notes
t	*z*	*p* (FWE)	x	y	z	
***p* <0.05 (FWE corrected)**						
7.21	5.61	0.00	42	-40	-25	Right vOT
6.96	5.48	0.00	-33	-40	-22	Left vOT
6.56	5.26	0.00	-45	-46	-19	Left vOT
6.13	5.02	0.01	39	-49	-22	Right vOT
6.07	4.98	0.01	-33	-85	-1	Left posterior middle occipital gyrus
5.97	4.92	0.01	-24	-91	-1	Left posterior middle occipital gyrus
5.89	4.87	0.01	33	-85	2	Right posterior middle occipital gyrus
5.72	4.78	0.02	42	-79	-7	Right amygdala
***p* <0.001, *k* >10 (uncorrected)**						
7.21	5.61	0.00	42	-40	-25	Right vOT
6.96	5.48	0.00	-33	-40	-22	Left vOT
6.56	5.26	0.00	-45	-46	-19	Left vOT
6.13	5.02	0.01	39	-49	-22	Right vOT
6.07	4.98	0.01	-33	-85	-1	Left posterior middle occipital gyrus
5.89	4.87	0.01	33	-85	2	Right posterior middle occipital gyrus
5.14	4.41	0.09	-33	14	-37	Left temporal pole
4.11	3.69	0.67	-24	-1	-22	Left amygdala
4.02	3.62	0.74	21	5	-19	Right anterior parahippocampal gyrus
3.91	3.54	0.82	-24	8	-19	Left Pars Orbitalis (BA 47)
3.80	3.45	0.89	-3	-34	-31	Left midbrain
3.74	3.41	0.91	-39	32	14	Left Pars Triangularis (BA 45) - within lFS

11.2 Picture Describing

Picture Describing (>rest)						
Peak values			MNI co-ordinates			Localisation notes
t	*z*	*p* (FWE)	x	y	z	
***p* <0.05 (FWE corrected)**						
8.79	5.83	0.00	-45	-10	32	Left precentral gyrus
7.57	5.36	0.00	-30	-91	-7	Left inferior occipital gyrus
7.50	5.33	0.00	48	-7	38	Right precentral gyrus
7.36	5.27	0.00	39	-55	-10	Right vOT
7.09	5.16	0.00	-54	-10	50	Left precentral gyrus
6.73	5.00	0.01	-45	-79	-1	Left middle occipital gyrus
6.31	4.80	0.02	-39	-34	-19	Left vOT
6.28	4.78	0.02	54	-1	47	Right precentral gyrus
6.11	4.70	0.03	36	-85	5	Right middle occipital gyrus
6.09	4.69	0.03	-30	5	-34	Left middle temporal pole
5.89	4.59	0.05	-33	-7	-28	Left anterior fusiform gyrus
5.89	4.59	0.05	3	14	62	Right SMA
***p* <0.001, *k* >10 (uncorrected)**						
8.79	5.83	0.00	-45	-10	32	Left precentral gyrus
7.57	5.36	0.00	-30	-91	-7	Left inferior occipital gyrus
7.50	5.33	0.00	48	-7	38	Right precentral gyrus
7.36	5.27	0.00	39	-55	-10	Right vOT
7.09	5.16	0.00	-54	-10	50	Left precentral gyrus
6.73	5.00	0.01	-45	-79	-1	Left middle occipital gyrus
6.31	4.80	0.02	-39	-34	-19	Left vOT
6.28	4.78	0.02	54	-1	47	Right precentral gyrus
6.11	4.70	0.03	36	-85	5	Right middle occipital gyrus
5.89	4.59	0.05	3	14	62	Right SMA
4.83	4.00	0.33	66	-1	20	Right precentral gyrus
4.78	3.97	0.36	0	2	65	Left SMA
4.28	3.65	0.70	-3	38	56	Left superior medial frontal gyrus

Colour Naming with sentences (>rest)						
Peak values			MNI co-ordinates			Localisation notes
t	*z*	*p* (FWE)	x	y	z	
***p* <0.05 (FWE corrected)**						
9.02	5.90	0.00	-45	-16	35	Left precentral gyrus
8.41	5.69	0.00	-21	-100	-1	Left middle occipital gyrus
7.83	5.46	0.00	24	-94	-1	Right inferior occipital gyrus
6.25	4.77	0.02	-54	-7	23	Left precentral gyrus
6.01	4.65	0.04	-63	-4	20	Left precentral gyrus
***p* <0.001, *k* >10 (uncorrected)**						
9.02	5.90	0.00	-45	-16	35	Left precentral gyrus
8.41	5.69	0.00	-21	-100	-1	Left middle occpital gyrus
7.83	5.46	0.00	24	-94	-1	Right inferior occipital gyrus
7.20	5.21	0.00	-21	-61	-25	Left cerebellum
6.25	4.77	0.02	-54	-7	23	Left precentral gyrus
6.01	4.65	0.04	-63	-4	20	Left precentral gyrus
5.85	4.57	0.05	48	-13	44	Right precentral gyrus
5.64	4.46	0.08	39	-13	44	Right precentral gyrus
5.33	4.29	0.15	9	2	59	Right SMA
5.28	4.26	0.16	-9	-7	56	Left SMA
4.98	4.09	0.27	3	2	68	Right SMA
4.60	3.86	0.50	54	-7	23	Right precentral gyrus
4.52	3.80	0.56	21	-10	-13	Right hippocampus
4.22	3.62	0.76	30	-7	-16	Right hippocampus
4.17	3.58	0.80	27	-4	-25	Right hippocampus
4.11	3.54	0.83	30	-64	35	Right middle occipital gyrus
4.10	3.53	0.84	36	-22	-13	Right hippocampus
4.09	3.53	0.84	36	-37	-13	Right vOT
4.07	3.51	0.85	33	-34	-4	Right hippocampus

Semantic and syntactic processing during sentence generation (PD>CNs)						
Peak values			MNI co-ordinates			Localisation notes
t	*z*	*p* (FWE)	x	y	z	
***p* <0.05 (FWE corrected)**						
6.50	4.89	0.02	-39	-40	-16	Left vOT
***p* <0.001, *k* >10 (uncorrected)**						
6.50	4.89	0.02	-39	-40	-16	Left vOT
4.91	4.04	0.34	-39	-34	-25	Left vOT
4.45	3.76	0.64	-42	-28	-19	Left vOT
4.71	3.92	0.46	-45	-76	-1	Left middle occipital gyrus
4.34	3.69	0.71	45	-34	-22	Right vOT
3.86	3.37	0.95	36	-34	-28	Right vOT
3.91	3.41	0.94	39	-49	-16	Right vOT

11.3 Listen and Name

Listen and Name (>rest)						
Peak values			MNI co-ordinates			Localisation notes
t	*z*	*p* (FWE)	x	y	z	
***p* <0.05 (FWE corrected)**						
11.30	7.29	0.00	54	-10	-1	Right STG (middle)
10.96	7.17	0.00	54	-19	2	Right STG (mid-posterior)
10.93	7.16	0.00	60	-25	5	Right STG (posterior)
10.47	7.00	0.00	-60	-13	5	Left STG (mid-posterior)
9.32	6.57	0.00	-60	-25	5	Left STG (posterior)
8.51	6.23	0.00	-36	-4	-31	Left parahippocampal gyrus / amygdala
8.21	6.09	0.00	3	5	65	SMA (BA 6)
8.10	6.04	0.00	-57	8	-10	Left temporal pole
8.09	6.03	0.00	-48	-10	41	Left precentral gyrus (dorsal)
7.68	5.84	0.00	-54	-4	20	Left precentral gyrus (ventral)
6.96	5.48	0.00	33	-1	-31	Right anterior fusiform gyrus/parahippocampal gyrus
6.91	5.45	0.00	48	-7	35	Right precentral gyrus (dorsal)
6.90	5.44	0.00	-51	32	2	Left Pars Triangularis (BA 45)
6.79	5.39	0.00	24	-1	-22	Right amygdala
6.52	5.24	0.00	-9	-25	-13	Left midbrain
6.41	5.18	0.00	-24	-10	-10	Left amygdala
6.21	5.07	0.00	30	8	-31	Right parahippocampal gyrus / amygdala
6.00	4.94	0.01	-45	17	-28	Left temporal pole
5.99	4.94	0.01	51	-4	26	Right precentral gyrus
5.72	4.78	0.02	-39	29	-1	Left Pars Triangularis (BA 45)
5.44	4.60	0.03	33	-37	-1	Right hippocampus (near trigone of lateral ventricle)
5.35	4.54	0.04	-3	29	56	Medial SFG (BA 8)
***p* <0.001, *k* >10 (uncorrected)**						
11.30	7.29	0.00	54	-10	-1	Right STG (middle)
10.96	7.17	0.00	54	-19	2	Right STG (mid-posterior)
10.93	7.16	0.00	60	-25	5	Right STG (posterior)
8.21	6.09	0.00	3	5	65	SMA
6.90	5.44	0.00	-51	32	2	Left Pars Triangularis
5.72	4.78	0.02	-39	29	-1	Left Pars Triangularis
5.44	4.60	0.03	33	-37	-1	Right hippocampus
5.35	4.54	0.04	-3	29	56	Frontal medial cortex (anterior to SMA)
4.97	4.29	0.11	-42	26	14	Left Pars Triangularis
4.22	3.77	0.50	0	41	50	Frontal superior medial cortex
4.16	3.72	0.55	27	-34	5	Right hippocampus/thalamus
3.76	3.42	0.85	-42	-43	-19	Left vOT

Alien Game (>rest)						
Peak values			MNI co-ordinates			Localisation notes
t	*z*	*p* (FWE)	x	y	z	
***p* <0.05 (FWE corrected)**						
8.69	6.30	0.00	60	-16	-1	Right STG (middle)
8.29	6.13	0.00	66	-22	2	Right STG (middle)
8.27	6.12	0.00	54	-28	-1	Right STG (posterior)
7.43	5.72	0.00	-66	-16	5	Left STG (middle)
7.15	5.58	0.00	-54	-25	8	Left STG (middle)
6.78	5.38	0.00	-57	-34	11	Left STG (posterior)
6.39	5.17	0.00	51	-7	44	Right precentral gyrus
5.85	4.85	0.01	63	5	-7	Right temporal pole
5.78	4.81	0.02	6	5	62	SMA
5.59	4.69	0.03	9	-1	62	SMA
***p* <0.001, *k* >10 (uncorrected)**						
8.69	6.30	0.00	60	-16	-1	Right STG (middle)
8.29	6.13	0.00	66	-22	2	Right STG (middle)
8.27	6.12	0.00	54	-28	-1	Right STG (posterior)
7.43	5.72	0.00	-66	-16	5	Left STG
7.15	5.58	0.00	-54	-25	8	Left STG (middle)
6.78	5.38	0.00	-57	-34	11	Left STG (posterior)
5.78	4.81	0.02	6	5	62	SMA
4.86	4.22	0.16	6	17	50	SMA
4.14	3.71	0.62	-6	8	53	SMA
5.03	4.33	0.11	-48	-10	47	Right precentral gyrus (dorsal)
4.98	4.30	0.12	-45	-13	38	Left precentral gyrus (dorsal)
5.00	4.31	0.11	24	-13	-10	Right globus pallidus
4.26	3.80	0.51	15	-4	-7	Right globus pallidus
4.20	3.75	0.57	12	-16	-16	Right midbrain
4.35	3.86	0.45	-24	-13	-7	Right lentiform nucleus
4.16	3.72	0.60	-33	17	-4	Left insula

Semantic and syntactic processing from phonological input (LN>AG)						
Peak values			MNI co-ordinates			Localisation notes
t	*z*	*p* (FWE)	x	y	z	
***p* <0.05 (FWE corrected)**						
6.56	5.26	0.00	21	-4	-19	Right hippocampus
6.11	5.01	0.01	-36	-10	-28	Left anterior fusiform gyrus
5.99	4.94	0.01	-45	-49	-16	Left vOT
5.64	4.72	0.02	-42	29	17	Left Pars Triangularis (BA 45)
5.57	4.68	0.03	-36	-40	-22	Left vOT
5.52	4.65	0.03	-60	-1	-10	Left anterior STS/STG
5.51	4.64	0.03	12	-4	-22	Right anterior parahippocampal gyrus / amygdala
5.37	4.56	0.05	39	-34	-19	Right vOT
***p* <0.001, *k* >10 (uncorrected)**						
6.56	5.26	0.00	21	-4	-19	Right hippocampus
6.11	5.01	0.01	-36	-10	-28	Left anterior fusiform gyurs
5.99	4.94	0.01	-45	-49	-16	Left vOT
5.64	4.72	0.02	-42	29	17	Left Pars Triangularis (BA 45)
5.57	4.68	0.03	-36	-40	-22	Left vOT
5.52	4.65	0.03	-60	-1	-10	Left anterior STG
5.51	4.64	0.03	12	-4	-22	Right parahippocampal gyrus
5.37	4.56	0.05	39	-34	-19	Right vOT
5.15	4.41	0.08	-9	-34	8	Left thalamus
4.87	4.23	0.15	21	-37	2	Right hippocampus (tail)
4.76	4.15	0.20	66	-1	2	Right anterior STG
4.75	4.14	0.20	33	8	-22	Right temporal pole
4.55	4.01	0.30	-9	-58	2	Left lingual gyrus
4.36	3.87	0.44	-3	35	56	Medial SFG (BA 8)
4.33	3.85	0.46	-54	35	2	Left Pars Triangularis (BA 45)
4.30	3.83	0.48	-39	11	26	Left Pars Triangularis (BA 45)
4.26	3.80	0.51	12	-61	11	Right calcerine sulcus
4.18	3.74	0.58	-27	23	-22	Left Pars Orbitalis (BA 47)
4.10	3.68	0.65	18	-67	11	Right calcerine sulcus
3.93	3.56	0.78	51	17	-25	Right temporal pole
3.87	3.50	0.83	0	-88	29	Left cuneus
3.84	3.48	0.84	57	14	-16	Right temporal pole
3.81	3.46	0.86	-42	29	2	Left Pars Triangularis (BA 45)
3.81	3.46	0.86	18	-73	2	Right lingual gyrus
3.51	3.23	0.98	-6	-19	8	Left thalamus

11.4 Read and Name

Read and Name (>rest)						
Peak values			MNI co-ordinates			Localisation notes
t	*z*	*p* (FWE)	x	y	z	
***p* <0.05 (FWE corrected)**						
12.12	6.13	0.00	-24	-91	-10	Left lingual gyrus (inferior occipital lobe)
10.48	5.77	0.00	27	-94	-7	Right lingual gyrus (inferior occipital lobe)
9.64	5.56	0.00	-42	20	20	Left Pars Triangualris (BA 45)
9.57	5.54	0.00	-51	-1	50	Left precentral gyrus (dorsal)
7.58	4.95	0.02	-42	-46	-13	Left vOT (approx.visual word form area)
***p* <0.001, *k* >10 (uncorrected)**						
12.12	6.13	0.00	-24	-91	-10	Left lingual gyrus (inferior occipital lobe)
10.48	5.77	0.00	27	-94	-7	Right lingual gyrus (inferior occipital lobe)
9.64	5.56	0.00	-42	20	20	Left Pars Triangualris (BA 47)
9.57	5.54	0.00	-51	-1	50	Left precentral gyrus (dorsal)
7.58	4.95	0.02	-42	-46	-13	Left vOT (approx.visual word form area)
6.62	4.59	0.06	-33	-4	-31	Left anterior fusiform gyrus
6.44	4.52	0.08	-39	5	26	Left Pars Opercularis/Triangularis (deep in IFS)
6.22	4.43	0.10	-45	-58	-13	Left posterior inferior temporal gyrus
6.21	4.43	0.11	33	-16	-13	Right anterior fusiform gyrus/hippocampus
5.77	4.24	0.20	51	-7	32	Right precentral gyrus
5.65	4.18	0.23	-51	-37	2	Left posterior MTG
5.65	4.18	0.23	-60	-37	2	Left posterior MTG
5.46	4.10	0.29	-6	11	53	SMA
5.11	3.92	0.45	30	2	-31	Right parahippocmapal gyrus
4.86	3.80	0.59	-33	-7	-22	Left anterior fusiform gyrus/hippocampus
4.81	3.77	0.62	57	-1	47	Right precentral gyrus (dorsal)
4.80	3.76	0.63	36	-7	-28	Right anterior fusiform gyrus
4.74	3.74	0.66	48	-1	56	Right precentral gyrus/middle frontal gyrus
4.71	3.72	0.67	12	-19	-7	Right midbrain
4.42	3.56	0.83	9	-10	-7	Right midbrain
4.01	3.32	0.96	-36	-34	-19	Left vOT

Face Finder (>rest)						
Peak values			MNI co-ordinates			Localisation notes
t	*z*	*p* (FWE)	x	y	z	
***p* <0.05 (FWE corrected)**						
7.72	4.99	0.01	33	-88	-7	Right posterior inferior occipital gyrus
7.46	4.90	0.02	24	-97	-13	Right lingual gyrus
7.29	4.84	0.03	30	-58	44	Right angular gyrus
7.09	4.77	0.03	-24	-64	41	Left angular gyrus
6.89	4.70	0.04	-30	-94	-10	Left posterior inferior occipital gyrus
***p* <0.001, *k* >10 (uncorrected)**						
7.72	4.99	0.01	33	-88	-7	Right posterior inferior occipital gyrus
7.46	4.90	0.02	24	-97	-13	Right lingual gyrus
5.08	3.91	0.51	42	-76	-7	Right posterior inferior occipital gyrus
7.29	4.84	0.03	30	-58	44	Right angular gyrus
6.76	4.65	0.05	30	-64	53	Right SPL
6.58	4.58	0.07	30	-64	35	Right middle/superior occipital gyrus/parietal lobe
7.09	4.77	0.03	-24	-64	41	Left angular gyrus
4.62	3.67	0.77	-30	-49	38	Left SMG
4.24	3.46	0.93	-30	-67	56	Left SPL
6.89	4.70	0.04	-30	-94	-10	Left posterior inferior occipital gyrus
5.59	4.16	0.27	-39	-67	-7	Left posterior inferior/middle occipital gyrus
4.92	3.83	0.60	-45	-70	-19	Left posterior FG
6.13	4.39	0.13	-45	-10	41	Left precentral gyrus
5.76	4.23	0.22	-27	8	11	Left putamen
4.76	3.74	0.69	-36	23	11	Left anterior insula (on border of BA 45)
4.21	3.43	0.94	-42	11	11	Left anterior operculum
5.36	4.04	0.37	3	14	56	Right SMA
4.93	3.83	0.59	0	-4	62	Left SMA
4.66	3.69	0.75	-39	44	8	Left lFS (BA45/MFG)
4.51	3.61	0.82	54	-7	35	Right premotor cortex
4.51	3.61	0.82	39	-16	38	Right precentral
4.41	3.55	0.87	-33	8	29	Right precentral gyrus
4.00	3.31	0.98	-39	2	26	Left Pars Opercularis (medial, on border with BA 4)
4.36	3.52	0.89	45	5	26	Right Pars Opercularis (medial, on border with BA 4)
3.94	3.27	0.98	45	23	32	Right MFG (just dorsal of BA 44/45)
4.28	3.48	0.92	42	-1	59	Right dorsal precentral gyrus
4.23	3.45	0.93	51	-49	-10	Right OT
3.72	3.14	1.00	54	-43	-16	Right OT
4.20	3.43	0.94	9	-22	-13	Right midbrain

Semantic and syntactic processing from orthographic input (RN>FF)						
Peak values			MNI co-ordinates			Localisation notes
t	*z*	*p* (FWE)	x	y	z	
***p* <0.001, *k* >10 (uncorrected)**						
6.88	4.69	0.06	6	-67	8	Right calcerine gyrus
6.18	4.42	0.15	-15	-73	11	Left calcerine gyrus
5.30	4.02	0.47	-21	-64	5	Left calcerine gyrus
5.29	4.01	0.48	-9	-94	-10	Left lingual gyrus
4.96	3.85	0.66	-15	-85	-4	Left lingual gyrus
4.44	3.57	0.91	3	-94	-7	Right lingual gyrus
5.16	3.95	0.55	-33	11	-34	Left temporal pole
3.91	3.26	1.00	-33	2	-31	Left temporal pole
5.03	3.89	0.62	57	-28	5	Right posterior STG
4.90	3.82	0.70	-54	35	-1	Left Pars Orbitalis (BA 47)
4.45	3.57	0.91	-48	26	-1	Left Pars Triangularis (BA 45)
4.85	3.79	0.73	-42	-46	-19	Left vOT

11.5 Semantic processing regions

Naming from pictures, orthography and phonology							
Peak values				MNI co-ordinates			Localisation notes
t	*z*	*p* (FWE)	*p* (unc)	x	y	z	
***p* <0.001, *k* >10 (uncorrected)**							
3.63	3.51	0.80	0.000	-42	-49	-22	Left vOT
3.43	3.33	0.93	0.000	30	-1	-34	Right anterior fusiform
3.42	3.32	0.94	0.000	-33	-34	-22	Left vOT (medial)
3.24	3.15	0.99	0.001	-33	14	-31	Left temporal pole
3.22	3.14	0.99	0.001	-36	17	-34	Left temporal pole
***p* <0.005, *k* >10 (uncorrected)**							
3.63	3.51	0.80	0.000	-42	-49	-22	Left vOT
3.43	3.33	0.93	0.000	30	-1	-34	Right anterior fusiform
3.42	3.32	0.94	0.000	-33	-34	-22	Left vOT (medial)
3.24	3.15	0.99	0.001	-33	14	-31	Left temporal pole
2.93	2.87	1.00	0.002	24	2	-22	Right anterior fusiform
2.92	2.86	1.00	0.002	-33	-4	-34	Left anterior fusiform gyrus
2.74	2.69	1.00	0.004	-42	-31	-31	Left vOT
2.65	2.60	1.00	0.005	-15	-7	-22	Left entorhinal cortex
2.65	2.60	1.00	0.005	-54	35	2	Left Pars Triangularis (BA 45)

12. GROUP MAPS RESULTS: AGE-RELATED CHANGES

Montreal Neurological Institute (MNI) co-ordinates for significant activation peaks (>8mm apart) are listed for each contrast performed in Chapter 8, along with *t*, *z*, and *p* values, and notes on the approximate localisation of activation based on the Automatic Anatomic Labels (Tzourio-Mazoyer et al., 2002) provided by the xjview toolbox (http://www.alivelearn.net/xjview8/). Peaks are listed according to activation strength (*t* value), from strongest to weakest, at two statistical thresholds (labelled).

12.1 Age-related changes in the auditory comprehension network (LN>rest)

Age-related increases in the auditory comprehension network (LN>rest)						
Peak values			MNI co-ordinates			Localisation notes
t	*z*	*p* (FWE)	x	y	z	
***p* <0.05 (FWE corrected)**						
7.15	5.55	0.00	27	-25	-19	Right parahippocampal gyrus (mid)
6.23	5.05	0.01	9	-13	-16	Right midbrain
5.45	4.59	0.04	3	-19	-16	Right midbrain
5.89	4.86	0.01	51	-61	-16	Right vOT
5.47	4.60	0.04	-6	-13	-13	Left midbrain
5.39	4.55	0.04	36	-34	-22	Right vOT
5.37	4.54	0.05	-36	32	-7	Left BA 47
***p* <0.00**						
7.15	5.55	0.00	27	-25	-19	Right parahippocampal gyrus (mid)
6.23	5.05	0.01	9	-13	-16	Right midbrain
5.89	4.86	0.01	51	-61	-16	Right vOT
5.37	4.54	0.05	-36	32	-7	Left BA 47
4.83	4.18	0.17	-18	-34	-1	Left hippocampus (tail) - BA 27
4.81	4.17	0.18	0	14	53	SMA
4.67	4.08	0.24	-48	2	47	Left MFG
4.62	4.04	0.26	48	-1	56	Right MFG
4.62	4.04	0.26	12	14	35	Right cingulate gyrus
4.56	4.00	0.30	60	-1	41	Right precentral gyrus
4.41	3.89	0.40	48	-7	-28	Right anterior ITS
4.40	3.89	0.40	63	-19	-7	Right STS (mid)
4.35	3.85	0.43	36	17	32	Right Pars Opercularis/DLPFC
4.35	3.85	0.44	60	-1	5	Right Heschl's gyrus
4.17	3.72	0.58	42	35	-13	Right Pars Orbitalis (BA 47)
4.16	3.71	0.59	-6	32	53	Medial superior frontal gyrus
4.11	3.67	0.63	60	-10	-10	Right STS (mid)
4.09	3.66	0.65	-48	17	38	Left MFG
3.98	3.58	0.74	-51	-4	-4	Left STG (mid)
3.94	3.55	0.77	-15	-91	-10	Left lingual gyrus
3.87	3.50	0.81	-54	-10	-10	Left STS (mid)
3.86	3.49	0.82	-51	-7	47	Left precentral gyrus
3.84	3.48	0.83	51	32	26	Right DLPFC
3.76	3.42	0.88	39	59	2	Right MFG (anterior)
3.74	3.40	0.89	39	56	11	Right MFG (anterior)
3.65	3.33	0.93	-27	23	-4	Left insula (anterior)
3.64	3.32	0.94	-33	17	-10	Left insula (anterior)
3.52	3.22	0.97	-27	-85	-19	Left lingual gyrus
3.43	3.16	0.98	42	50	17	Right DLPFC

12.2 Age-related changes in the semantic processing network (LN>AG)

Age-related increases in the semantic processing regions (LN>AG)						
Peak values			MNI co-ordinates			Localisation notes
t	*z*	*p* (FWE)	x	y	z	
***p* <0.05 (FWE corrected): LN>AG mask**						
6.77	5.35	0.00	39	-37	-19	Right vOT (anterior)
***p* <0.001, *k* >10 (uncorrected): LN>AG mask**						
6.77	5.35	0.00	39	-37	-19	Right vOT (anterior)
4.16	3.72	0.62	-48	-67	-10	Left vOT
4.10	3.67	0.67	-36	-37	-19	Left vOT (anterior)
3.55	3.25	0.97	-51	-61	-1	Left vOT

13. A NEW VOXEL-BASED ARTEFACT REPAIR METHOD

This abstract was written by Tim Tierney for the International Society for Magnetic Resonance in Medicine Movement Correction in MRI workshop (2014), and represents development of a new movement artefact repair method, trialled on a subset of data from this thesis.

A Biophysical Model for Retrospective Motion Correction in fMRI and a Comparison of Current Methodologies

PURPOSE: Currently motion correction approaches applied to the 2D gradient echo EPI typically used to obtain fMRI data do not account for rapid changes in head position that occur at a timescale less than the volume repetition rate (TR). Usually volume to volume spatial realignment is used and the transformation parameters are employed as a model of head motion1. In ascending slice-ordered sequences this type of movement can result in slice specific signal drop-out > 30%. This is due to parts of a slice being excited twice in rapid succession with associated spin history effects. While methods exist to correct for these problems, many rely on modelling derivatives of the transformation parameters. However as the resulting regressors are sampled at 1/TR they cannot account for movement effects between volumes. A different approach is to use robust least squares to de-weight entire volumes if their residual variance is high2. However in the case of slice specific dropouts the assumption of no spatial structure to the noise is violated and the method is also limited. Therefore we propose to develop a biophysical model that identifies spurious signal variations at every voxel. The model then corrects the time series in order to address the limitations of these other methods

METHODS:

Method Theory: To identify spurious variation a threshold needs to be set in terms of percentage signal intensity change. To estimate the percentage signal intensity change relative to baseline we assume that the Blood Oxygen Level Dependent signal can be modelled according to the following equation: $S = Smax*e^{-TE/T2*}$, where S = BOLD Signal Intensity, Smax =100, TE= echo time and $T2* = 1/R2 + 1/R2'$, R2 = spin- spin relaxation rate and R2' = relaxation due to magnetic field inhomogeneities. Empirical and theoretical models allow for the estimation of R2 and R2' 3,4. To make a robust threshold we can estimate $Smax*e^{-TE/T2} - Smax*e^{-TE/T2*}$, where T2 =1/R2. The resulting threshold now represents the percentage deviation from baseline that would be observed in the BOLD signal should a voxel be completely depleted of deoxyhaemoglobin. It is therefore plausible that signals outside this range represent non-physiological signals and artefacts. Once these non-physiological signals are identified they can be spline interpolated to correct them.

The primary limitation of this model is that it will break down in areas where the noise is exceptionally high such as at the edge of the brain and in the veins and arteries. However we can segment these areas from the BOLD images by using the Expectation Maximisation (EM) algorithm on the median/(median absolute deviation (MAD)) image. Once these noisy areas are segmented the first six principal components are extracted. These regressors contain information on pulsatile motion (from the arteries) and of subject head movement (Figure 1). They can be included as effects of no interest in one's regression analysis in order to improve one's model of motion.

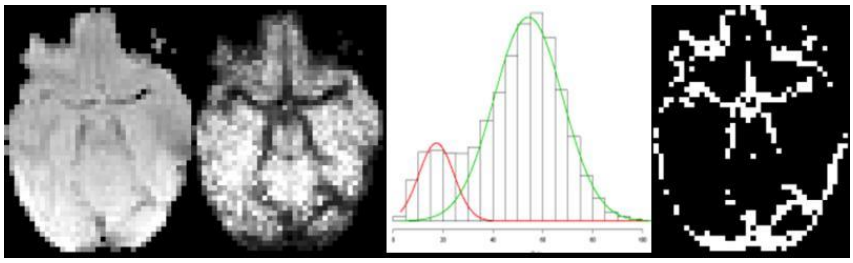


FIGURE 1: Step 1: The median image is calculated. Step 2: Divide by the MAD image. Step 3: Segment using ESM. Step 4: Extract principal components from the resulting mask. A lower axial slice is shown to demonstrate the ability to observe both the Circle of Willis and the Transverse sinus.

Retrospective Correction Techniques Comparison: A number of other retrospective motion correction methods were also examined for the purpose of comparison. The other methods included were: Robust Weighted Least Squares (RWLS)2, Motion Fingerprint (MF)5, Realignment Parameter Expansion(RPE)1,6 and simply including the realignment parameters as effects of no interest in the GLM.

Data Collection: Data was collected from 42 children aged 6-18 while performing a listening comprehension task. This task was chosen as there was a strong a priori hypothesis about what constituted the ground truth (Bilateral superior temporal gyrus, left inferior frontal, bilateral cerebellum, left temporal pole, bilateral supplementary motor area) and the data was severely corrupted by motion artefacts. The data was obtained using a 1.5T Siemens Avanto scanner, TR = 2.16s, TE=30 ms, FOV=210mm, flip angle=75 degrees, number of slices=30, slice thickness= 3mm, slice gap=1mm. There were 212 volumes acquired using continuous image acquisition, covering the brain with resolution 3.3 x 3.3 x 4mm.

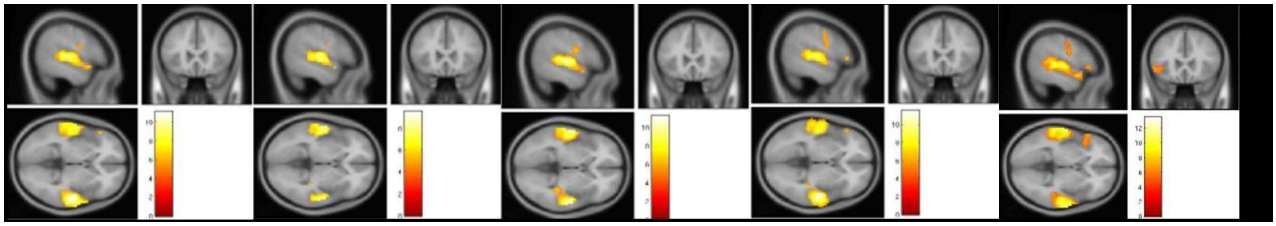


FIGURE 2: Comparison of retrospective motion correction techniques. The t maps are overlaid on a normalised T1 and are thresholded at $p<0.05$ (FWE corrected).

RESULTS: Figure 2 displays a comparison of the 5 different methods examined. It was found that both RPE and MF had a negative impact on the group map as far fewer voxels were identified in the a priori specified regions of interest compared to the standard GLM, most notably below in the bilateral superior temporal gyri and the lack of activation in the left inferior frontal gyrus. RWLS recovers some activation in the inferior frontal gyrus but the extension of the superior temporal gyrus to the temporal pole is lost. Furthermore the activations do not follow the shape of the cortex and present with a number of false positives in the white matter. The method which uses the biophysical model performs qualitatively best as it recovers a substantial amount of the inferior frontal gyrus which none of the other methods do. This method also strengthens the activations identified using the standard GLM. All clusters in the Biophysical approach display larger local maxima than all the other methods in the a priori specified regions of interest.

To assess the impact of the model quantitatively we calculated mutual information between each individual t-map and the group map. The resulting method would therefore represent how the method changes the inter-subject variability in activations. Mutual Information (MI) was calculated between the group t-maps of the corrected and uncorrected datasets and their respective first level t-maps. The effect size, r, was calculated 7. The mean of the difference was .118 (95% CI = .095, .141), t (41)= 10.311, p < .05, r = .845.

DISCUSSION: The method proposed results in a substantial qualitative (recovery of inferior frontal gyrus) and quantitative improvement (increased local maxima and lower individual variability). However it must be noted that while a comparison of methodologies is presented this method is not directly competitive with these other methods. It can be used in addition to the others reviewed. The primary limitation of this method is the direct imputation of values due to interpolation. For future work it may be worthwhile to consider multiple imputation as opposed to spline interpolation.

CONCLUSIONS: The proposed method can improve severely corrupted datasets. This has direct clinical relevance in the pre-surgical evaluation of patients who need to undergo either language or motor mapping such as those with focal epilepsy.

ABSTRACT REFERENCES:

1. Friston, K. J., Williams, S., Howard, R., Frackowiak, R. S., & Turner, R. (1996). Movement-related effects in fMRI time-series. Magn. Reson. Med, 35(3), 346-55.
2. Diedrichsen, J., & Shadmehr, R. (2005). Detecting and adjusting for artefacts in fMRI time series data. Neuroimage, 27(3), 624-634
3. Uludağ,K. Müller-Bierl,B. Uğurbil,K.(2009) An integrative model for neuronal activity-induced signal changes for gradient and spin echo functional imaging, NeuroImage, 48, 1.
4. Yablonsky, A., Haacke, M., (1994). Theory of NMR signal behavior in magnetically inhomogeneous tissues: the static dephasing regime. Magn. Reson. Med. 32,(15), 749–763
5. Wilke,M.(2012). An alternative approach towards assessing and accounting for individual motion in fMRI timeseries, NeuroImage, 59, (3), 2062-2072
6. Lemieux, L., Salek-Haddadi, A., Lund, T. E., Laufs, H., & Carmichael, D. (2007). Modelling large motion events in fMRI studies of patients with epilepsy. Magnetic resonance imaging, 25(6), 894-901
7.Rosnow,L.,Rosenthal,R & Rubin,B.(2000).Contrasts and Correlations in effect size estimation.

Psychological Science,11,446-453

REFERENCES

Adank, P., & Devlin, J. T. (2010). On-line plasticity in spoken sentence comprehension: Adapting to time-compressed speech. *NeuroImage, 49*(1), 1124-1132.

Adhami, F., Liao, G., Morozov, Y. M., Schloemer, A., Schmithorst, V. J., Lorenz, J. N., . . . Kuan, C. Y. (2006). Cerebral ischemia-hypoxia induces intravascular coagulation and autophagy. *Am J Pathol, 169*(2), 566-583.

Adleman, N. E., Menon, V., Blasey, C. M., White, C. D., Warsofsky, I. S., Glover, G. H., & Reiss, A. L. (2002). A Developmental fMRI Study of the Stroop Color-Word Task. *NeuroImage, 16*(1), 61-75.

Adlington, R. L., Laws, K. R., & Gale, T. M. (2009). The Hatfield Image Test (HIT): a new picture test and norms for experimental and clinical use. *J Clin Exp Neuropsychol, 31*(6), 731-753.

Agosta, F., Galantucci, S., Canu, E., Cappa, S. F., Magnani, G., Franceschi, M., . . . Filippi, M. (2013). Disruption of structural connectivity along the dorsal and ventral language pathways in patients with nonfluent and semantic variant primary progressive aphasia: A DT MRI study and a literature review. *Brain & Language, in press*.

Ahmad, Z., Balsamo, L. M., Sachs, B. C., Xu, B., & Gaillard, W. D. (2003). Auditory comprehension of language in young children: neural networks identified with fMRI. *Neurology, 60*(10), 1598-1605.

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology, 8*(2), 71-82.

Andersson, J. L., Ashburner, J., & Friston, K. (2001). A global estimator unbiased by local changes. *NeuroImage, 13*(6 Pt 1), 1193-1206.

Anwander, A., Tittgemeyer, M., von Cramon, D. Y., Friederici, A. D., & Knosche, T. R. (2007). Connectivity-Based Parcellation of Broca's Area. *Cerebal Cortex, 17*(4), 816-825.

Argyropoulos, G. P., Tremblay, P., & Small, S. L. (2013). The neostriatum and response selection in overt sentence production: an fMRI study. *NeuroImage, 82*, 53-60.

Arora, J., Pugh, K., Westerveld, M., Spencer, S., Spencer, D. D., & Todd Constable, R. (2009). Language lateralization in epilepsy patients: fMRI validated with the Wada procedure. *Epilepsia, 50*(10), 2225-2241.

Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *NeuroImage, 38*(1), 95-113.

Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *NeuroImage, 26*(3), 839-851.

Awad, M., Warren, J. E., Scott, S. K., Turkheimer, F. E., & Wise, R. J. (2007). A common system for the comprehension and production of narrative speech. *J Neurosci, 27*(43), 11455-11464.

Baciu, M., Juphard, A., Cousin, E., & Bas, J. F. (2005). Evaluating fMRI methods for assessing hemispheric language dominance in healthy subjects. *Eur J Radiol, 55*(2), 209-218.

Baciu, M. V., Watson, J. M., Maccotta, L., McDermott, K. B., Buckner, R. L., Gilliam, F. G., & Ojemann, J. G. (2005). Evaluating functional MRI procedures for assessing hemispheric language dominance in neurosurgical patients. *Neuroradiology, 47*(11), 835-844.

Bailet, L. L., & Turk, W. R. (2000). The impact of childhood epilepsy on neurocognitive and behavioral performance: a prospective longitudinal study. *Epilepsia, 41*(4), 426-431.

Balsamo, L. M., Xu, B., & Gaillard, W. D. (2006). Language lateralization and the role of the fusiform gyrus in semantic processing in young children. *NeuroImage, 31*(3), 1306-1314.

Bandettini, P. A., Wong, E. C., Hinks, R. S., Tikofsky, R. S., & Hyde, J. S. (1992). Time course EPI of human brain function during task activation. *Magn Reson Med, 25*(2), 390-397.

Barch, D. M., Sabb, F. W., Carter, C. S., Braver, T. S., Noll, D. C., & Cohen, J. D. (1999). Overt verbal responding during fMRI scanning: empirical investigations of problems and potential solutions. *NeuroImage, 10*(6), 642-657.

Barry, C., Morrison, C. M., & Ellis, A. W. Naming the Snodgrass and Vanderwart Pictures: Effects of Age of Acquisition, Frequency, and Name Agreement. *Quarterly Journal of Experimental Psychology: Section A. Aug97, 50*(3), 560.

Bates, E. G., J.C. . (1997). ON THE INSEPARABILITY OF GRAMMAR AND THE LEXICON: EVIDENCE FROM ACQUISITION, APHASIA AND REAL-TIME PROCESSING. *Language and Cognitive Processes, 12*(5/6), 507-584.

Bauer, P. R., Reitsma, J. B., Houweling, B. M., Ferrier, C. H., & Ramsey, N. F. (2014). Can fMRI safely replace the Wada test for preoperative assessment of language lateralisation? A meta-analysis and systematic review. *J Neurol Neurosurg Psychiatry, 85*(5), 581-588.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Methodological, 57*(1), 289-300.

Bennett, C. M., Miller, M. B., & Wolford, G. L. (2009). Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: an argument for multiple comparisons correction. *NeuroImage, 47, Supplement 1*(0), S125.

Benson, R. R., FitzGerald, D. B., LeSueur, L. L., Kennedy, D. N., Kwong, K. K., Buchbinder, B. R., . . . Rosen, B. R. (1999). Language dominance determined by whole brain functional MRI in patients with brain lesions. *Neurology, 52*(4), 798-809.

Benson, R. R., Whalen, D. H., Richardson, M., Swainson, B., Clark, V. P., Lai, S., & Liberman, A. M. (2001). Parametrically dissociating speech and nonspeech perception in the brain using fMRI. *Brain Lang, 78*(3), 364-396.

Berg, A. T. (2011). Epilepsy, cognition, and behavior: The clinical picture. *Epilepsia, 52 Suppl 1*, 7-12.

Berg, A. T., Berkovic, S. F., Brodie, M. J., Buchhalter, J., Cross, J. H., van Emde Boas, W., . . . Scheffer, I. E. (2010). Revised terminology and concepts for organization of seizures and epilepsies: report of the ILAE Commission on Classification and Terminology, 2005-2009. *Epilepsia, 51*(4), 676-685.

Berl, M. M., Duke, E. S., Mayo, J., Rosenberger, L. R., Moore, E. N., VanMeter, J., . . . Gaillard, W. D. (2010). Functional anatomy of listening and reading comprehension during development. *Brain Lang, 114*(2), 115-125.

Berl, M. M., Mayo, J., Parks, E. N., Rosenberger, L. R., Vanmeter, J., Ratner, N. B., . . . Gaillard, W. D. (2014). Regional differences in the developmental trajectory of lateralization of the language network. *Human Brain Mapping, 35*(1), 270-284.

Berl, M. M., Zimmaro, L. A., Khan, O. I., Dustin, I., Ritzl, E., Duke, E. S., . . . Gaillard, W. D. (2013). Characterization of atypical language activation patterns in focal epilepsy. *Ann Neurol*.

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends Cogn Sci, 15*(11), 527-536.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb Cortex, 19*(12), 2767-2796.

Binder, J. R., Gross, W. L., Allendorfer, J. B., Bonilha, L., Chapin, J., Edwards, J. C., . . . Weaver, K. E. (2011). Mapping anterior temporal lobe language areas with fMRI: a multicenter normative study. *NeuroImage, 54*(2), 1465-1475.

Binder, J. R., Swanson, S. J., Sabsevitz, D. S., Hammeke, T. A., Raghavan, M., & Mueller, W. M. (2010). A comparison of two fMRI methods for predicting verbal memory decline after left temporal lobectomy_Language lateralization versus hippocampal activation asymmetry. *Epilepsia, 51*(4), 618-626.

Binney, R. J., Parker, G. J. M., & Ralph, M. A. L. (2012). Convergent Connectivity and Graded Specialization in the Rostral Human Temporal Lobe as Revealed by Diffusion-Weighted Imaging Probabilistic Tractography. *J Cogn Neurosci, 24*(10), 1998-2014.

Bishop, D. V. (2013). Cerebral asymmetry and language development: cause, correlate, or consequence? *Science, 340*(6138), 1230531.

Bitan, T., Burman, D. D., Lu, D., Cone, N. E., Gitelman, D. R., Mesulam, M. M., & Booth, J. R. (2006). Weaker top-down modulation from the left inferior frontal gyrus in children. *NeuroImage, 33*(3), 991-998.

Bjornaes, H., Stabell, K., Henriksen, O., & Loyning, Y. (2001). The effects of refractory epilepsy on intellectual functioning in children and adults. A longitudinal study. *Seizure, 10*(4), 250-259.

Blakemore, S. J., & Choudhury, S. (2006). Development of the adolescent brain: implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry, 47*(3-4), 296-312.

Blank, S. C., Scott, S. K., Murphy, K., Warburton, E., & Wise, R. J. (2002). Speech production: Wernicke, Broca and beyond. *Brain, 125*(Pt 8), 1829-1838.

Blesser, B. (1972). Speech Perception under Conditions of Spectral Transformation .1. Phonetic Characteristics. *J Speech Hear Res, 15*(1), 5-&.

Bookheimer, S. (2002). Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Annu Rev Neurosci, 25*, 151-188.

Booth, J. R., Burman, D. D., Van Santen, F. W., Harasaki, Y., Gitelman, D. R., Parrish, T. B., & Marsel Mesulam, M. M. (2001). The development of specialized brain systems in reading and oral-language. *Child Neuropsychol, 7*(3), 119-141.

Brauer, J., Anwander, A., & Friederici, A. D. (2011). Neuroanatomical prerequisites for language functions in the maturing brain. *Cerebral Cortex, 21*(2), 459-466.

Brauer, J., Anwander, A., Perani, D., & Friederici, A. D. (2013). Dorsal and ventral pathways in language development. *Brain & Language, In press*.

Brauer, J., & Friederici, A. D. (2007). Functional neural networks of semantic and syntactic processes in the developing brain. *J Cogn Neurosci, 19*(10), 1609-1623.

Brett, M., Anton, J. L., Valabregue, R., & Poline, J. B. (2002). *Region of interest analysis using the MarsBar toolbox for SPM 99*. Paper presented at the 8th International Conferance on Functional Mapping of the Human Brain, Sendai, Japan. http://matthew.dynevor.org/_downloads/marsbar.pdf

The British National Corpus. (2007). from http://www.natcorp.ox.ac.uk/

Broca, P. (1865). Sur le siège de la faculté du langage articulé. *Bulletins de la Societe d'Anthropologie de Paris, 6*(1), 377-393.

Brodmann, K. (1909). *Brodmann's localisation in the human cortex* ((traduction anglaise par Laurence Garey of the German book) ed.). London, UK: Smith-Gordon.

Brown, T. T., Lugar, H. M., Coalson, R. S., Miezin, F. M., Petersen, S. E., & Schlaggar, B. L. (2005). Developmental changes in human cerebral functional organization for word generation. *Cerebral Cortex, 15*(3), 275-290.

Buchel, C., Price, C., & Friston, K. (1998). A multimodal language region in the ventral visual pathway. *Nature, 394*(6690), 274-277.

Burgund, E. D., Kang, H. C., Kelly, J. E., Buckner, R. L., Snyder, A. Z., Petersen, S. E., & Schlaggar, B. L. (2002). The feasibility of a common stereotactic space for children and adults in fMRI studies of development. *NeuroImage, 17*(1), 184-200.

Byars, A. W., Holland, S. K., Strawsburg, R. H., Bommer, W., Dunn, R. S., Schmithorst, V. J., & Plante, E. (2002). Practical aspects of conducting large-scale functional magnetic resonance imaging studies in children. *J Child Neurol, 17*(12), 885-890.

Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *J Cogn Neurosci, 12*(1), 1-47.

Cai, Q., Lavidor, M., Brysbaert, M., Paulignan, Y., & Nazir, T. A. (2008). Cerebral lateralization of frontal lobe language processes and lateralization of the posterior visual word processing system. *J Cogn Neurosci, 20*(4), 672-681.

Caplan, R., Siddarth, P., Vona, P., Stahl, L., Bailey, C., Gurbani, S., . . . Donald Shields, W. (2009). Language in pediatric epilepsy. *Epilepsia, 50*(11), 2397-2407.

Carroll, J. B., & White, M. N. (1973). Word frequency and age of acquisition as determiners of picture-naming latency. *Quarterly Journal of Experimental Psychology, 25*(1).

Casey, B. J., Galvan, A., & Hare, T. A. (2005). Changes in cerebral functional organization during cognitive development. *Curr Opin Neurobiol, 15*(2), 239-244.

Catani, M., Jones, D. K., & ffytche, D. H. (2005). Perisylvian language networks of the human brain. *Annals of Neurology, 57*(1), 8-16.

Catani, M., & Mesulam, M. (2008). The arcuate fasciculus and the disconnection theme in language and aphasia: history and current state. *Cortex, 44*(8), 953-961.

Cattell, J. M. (1886). *Psychometrische Untersuchungen.* Retrieved from http://www.uni-leipzig.de/~psycho/wundt/opera/cattell/psymtrik/PSYMETUI.htm

Chaudhary, U. J., Kokkinos, V., Rodionov, R., Carmichael, D. W., Vulliemoz, S., Thornton, R., . . . Lemieux, L. (2009). Implementation of Video-Eeg-Fmri to Investigate Epileptic Activity. *Epilepsia, 50*, 72-73.

Chou, T. L., Booth, J. R., Burman, D. D., Bitan, T., Bigio, J. D., Lu, D., & Cone, N. E. (2006). Developmental changes in the neural correlates of semantic processing. *NeuroImage, 29*(4), 1141-1149.

Christoff, K., Prabhakaran, V., Dorfman, J., Zhao, Z., Kroger, J. K., Holyoak, K. J., & Gabrieli, J. D. E. (2001). Rostrolateral Prefrontal Cortex Involvement in Relational Integration during Reasoning. *NeuroImage, 14*(5), 1136-1149.

Christoffels, I. K., Formisano, E., & Schiller, N. O. (2007). Neural correlates of verbal feedback processing: an fMRI study employing overt speech. *Human Brain Mapping, 28*(9), 868-879.

Chumbley, J., Worsley, K., Flandin, G., & Friston, K. (2010). Topological FDR for neuroimaging. *NeuroImage, 49*(4), 3057-3064.

Chumbley, J. R., & Friston, K. J. (2009). False discovery rate revisited: FDR and topological inference using Gaussian random fields. *NeuroImage, 44*(1), 62-70.

Clusmann, H., Kral, T., Gleissner, U., Sassen, R., Urbach, H., Blumcke, I., . . . Schramm, J. (2004). Analysis of different types of resection for pediatric patients with temporal lobe epilepsy. *Neurosurgery, 54*(4), 847-859; discussion 859-860.

Cohen, L., Lehericy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain, 125*(Pt 5), 1054-1069.

Collignon, A., Maes, F., Delaere, D., Vandermeulen, D., Suetens, P., & Marchal, G. (1995). Automated multi-modality image registration based on information theory. *Information Processing in Medical Imaging, 3*, 263-274.

Constable, R. T., Pugh, K. R., Berroya, E., Mencl, W. E., Westerveld, M., Ni, W. J., & Shankweiler, D. (2004). Sentence complexity and input modality effects in sentence comprehension: an fMRI study. *NeuroImage, 22*(1), 11-21.

Corina, D. P., Loudermilk, B. C., Detwiler, L., Martin, R. F., Brinkley, J. F., & Ojemann, G. (2010). Analysis of naming errors during cortical stimulation mapping: implications for models of language representation. *Brain Lang, 115*(2), 101-112.

Cormack, F., Cross, J. H., Isaacs, E., Harkness, W., Wright, I., Vargha-Khadem, F., & Baldeweg, T. (2007). The development of intellectual abilities in pediatric temporal lobe epilepsy. *Epilepsia, 48*(1), 201-204.

Crinion, J., & Price, C. J. (2005). Right anterior superior temporal activation predicts auditory sentence comprehension following aphasic stroke. *Brain, 128*(Pt 12), 2858-2871.

Croft, L. J., Baldeweg, T., Sepeta, L. S., Zimmaro, L. A., Berl, M. M., & Gaillard, W. D. (in press). Vulnerability of the ventral language network in children with focal epilepsy. *Brain*.

Croft, L. J., Rankin, P. M., Liegeois, F., Banks, T., Cross, J. H., Vargha-Khadem, F., & Baldeweg, T. (2013). To speak, or not to speak? The feasibility of imaging overt speech in children with epilepsy. *Epilepsy Res*.

Cycowicz, Y. M., Friedman, D., Rothstein, M., & Snodgrass, J. G. (1997). Picture naming by young children: norms for name agreement, familiarity, and visual complexity. *J Exp Child Psychol, 65*(2), 171-237.

Davies, K. G., Bell, B. D., Bush, A. J., Hermann, B. P., Dohan, F. C., Jr., & Jaap, A. S. (1998). Naming decline after left anterior temporal lobectomy correlates with pathological status of resected hippocampus. *Epilepsia, 39*(4), 407-419.

Davies, K. G., Maxwell, R. E., Jennum, P., Dhuna, A., Beniak, T. E., Destafney, E., . . . Fiol, M. E. (1994). Language function following subdural grid-directed temporal lobectomy. *Acta Neurol Scand, 90*(3), 201-206.

Davies, K. G., Risse, G. L., & Gates, J. R. (2005). Naming ability after tailored left temporal resection with extraoperative language mapping: increased risk of decline with later epilepsy onset age. *Epilepsy & Behavior: E&B, 7*(2), 273-278.

Davies, S., Heyman, I., & Goodman, R. (2003). A population survey of mental health problems in children with epilepsy. *Dev Med Child Neurol, 45*(5), 292-295.

de Bie, H. M., Boersma, M., Wattjes, M. P., Adriaanse, S., Vermeulen, R. J., Oostrom, K. J., . . . Delemarre-Van de Waal, H. A. (2010). Preparing children with a mock scanner training protocol results in high quality structural and functional MRI scans. *Eur J Pediatr, 169*(9), 1079-1085.

de Ribaupierre, S., Fohlen, M., Bulteau, C., Dorfmuller, G., Delalande, O., Dulac, O., . . . Hertz-Pannier, L. (2012). Presurgical language mapping in children with epilepsy: clinical usefulness of functional magnetic resonance imaging for the planning of cortical stimulation. *Epilepsia, 53*(1), 67-78.

DeLeon, J., Gottesman, R. F., Kleinman, J. T., Newhart, M., Davis, C., Heidler-Gary, J., . . . Hillis, A. E. (2007). Neural regions essential for distinct cognitive processes underlying picture naming. *Brain, 130*(Pt 5), 1408-1422.

Demonet, J. F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J. L., Wise, R., . . . Frackowiak, R. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain, 115 (Pt 6)*, 1753-1768.

Demonet, J. F., Price, C., Wise, R., & Frackowiak, R. S. (1994). Differential activation of right and left posterior sylvian regions by semantic and phonological tasks: a positron-emission tomography study in normal human subjects. *Neurosci Lett, 182*(1), 25-28.

Dhanjal, N. S., Handunnetthi, L., Patel, M. C., & Wise, R. J. (2008). Perceptual systems controlling speech production. *J Neurosci, 28*(40), 9969-9975.

Dichter, M. A., & Brodie, M. J. (1996). New antiepileptic drugs. *N Engl J Med, 334*(24), 1583-1590.

Diedrichsen, J., & Shadmehr, R. (2005). Detecting and adjusting for artifacts in fMRI time series data. *NeuroImage, 27*(3), 624-634.

Eichenbaum, H. (2000). A cortical-hippocampal system for declarative memory. *Nat Rev Neurosci, 1*(1), 41-50.

Elliott, C. D. (1990). *Differential ability scales: Introductory and technical handbook*. San Antonio: The Psychological Corp.

Etard, O., Mellet, E., Papathanassiou, D., Benali, K., Houde, O., Mazoyer, B., & Tzourio-Mazoyer, N. (2000). Picture naming without Broca's and Wernicke's area. *Neuroreport, 11*(3), 617-622.

Everts, R., Lidzba, K., Wilke, M., Kiefer, C., Mordasini, M., Schroth, G., . . . Steinlin, M. (2009). Strengthening of laterality of verbal and visuospatial functions during childhood and adolescence. *Human Brain Mapping, 30*(2), 473-483.

Farah, M. J., Mcmullen, P. A., & Meyer, M. M. (1991). Can Recognition of Living Things Be Selectively Impaired. *Neuropsychologia, 29*(2), 185-193.

Fernandez, G., Specht, K., Weis, S., Tendolkar, I., Reuber, M., Fell, J., . . . Elger, C. E. (2003). Intrasubject reproducibility of presurgical language lateralization and mapping using fMRI. *Neurology, 60*(6), 969-975.

Fiebach, C. J., Schlesewsky, M., Lohmann, G., von Cramon, D. Y., & Friederici, A. D. (2005). Revisiting the role of Broca's area in sentence processing: syntactic integration versus syntactic working memory. *Human Brain Mapping, 24*(2), 79-91.

Field, A. (2005). *Discovering Statistics Using SPSS* (Second Edition ed.). London, UK: SAGE Publications.

Fink, G. R., Halligan, P. W., Marshall, J. C., Frith, C. D., Frackowiak, R. S., & Dolan, R. J. (1996). Where in the brain does visual attention select the forest and the trees? *Nature, 382*(6592), 626-628.

FitzGerald, D. B., Cosgrove, G. R., Ronner, S., Jiang, H., Buchbinder, B. R., Belliveau, J. W., . . . Benson, R. R. (1997). Location of language in the cortex: a comparison between functional MR imaging and electrocortical stimulation. *AJNR. American Journal of Neuroradiology, 18*(8), 1529-1539.

Forman, S. D., Cohen Jd Fau - Fitzgerald, M., Fitzgerald M Fau - Eddy, W. F., Eddy Wf Fau - Mintun, M. A., Mintun Ma Fau - Noll, D. C., & Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magnetic Resonance in Medicine, 33*(0740-3194 (Print)), 636-647.

Forster, K. I., & Forster, J. C. (2010). DMDX: A Windows Display Program with Millisecond Accuracy. *Behavior Research Methods Instruments and Computers, Vol. 35, No. 1. (2003), pp. 116-124, 35*(1), 8.

Fox, P. T., & Raichle, M. E. (1986). Focal physiological uncoupling of cerebral blood flow and oxidative metabolism during somatosensory stimulation in human subjects. *Proceedings of the National Academy of Science USA, 83*, 1140-1144.

Fox, P. T., & Raichle, M. E. (1988). Nonoxidative glucose consumption during focal physiological neural activity. *Science, 241*, 462-464.

Frahm, J., Kruger, G., Merboldt, K. D., & Kleinschmidt, A. (1996). Dynamic uncoupling and recoupling of perfusion and oxidative metabolism during focal brain activation in man. *Magnetic Resonance in Medicine, 35*, 143-148.

Fraisse, P. (1969). Why is naming longer than reading? *Acta Psychologica, 30*, 96–103.

Frederiksen, J. R., & Kroll, J. F. (1976). Spelling and sound: Approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception and Performance, 2*(3), 361.

Frey, S., Campbell, J. S., Pike, G. B., & Petrides, M. (2008). Dissociating the human language pathways with high angular resolution diffusion fiber tractography. *Journal of Neuroscience, 28*(45), 11435-11444.

Friederici, A. D. (2006). Broca's area and the ventral premotor cortex in language: functional differentiation and specificity. *Cortex, 42*(4), 472-475.

Friederici, A. D. (2012). Language development and the ontogeny of the dorsal pathway. *Front Evol Neurosci, 4*, 3.

Friederici, A. D., Bahlmann, J., Heim, S., Schubotz, R. I., & Anwander, A. (2006). The brain differentiates human and non-human grammars: functional localization and structural connectivity. *Proceedings of the National Academy of Sciences USA, 103*(7), 2458-2463.

Friederici, A. D., Brauer, J., & Lohmann, G. (2011). Maturation of the language network: from inter- to intrahemispheric connectivities. *PLoS One, 6*(6), e20726.

Friston, K., Moran, R., & Seth, A. K. (2013). Analysing connectivity with Granger causality and dynamic causal modelling. *Curr Opin Neurobiol, 23*(2), 172-178.

Friston, K. J., Frith, C. D., Frackowiak, R. S., & Turner, R. (1995). Characterizing dynamic brain responses with fMRI: a multivariate approach. *NeuroImage, 2*(2), 166-172.

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *NeuroImage, 19*(4), 1273-1302.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D., & Frackowiak, R. S. J. (1995). Statistical Parametric Maps in Functional Imaging: A General Linear Approach. *Human Brain Mapping, 2*, 189-210.

Friston, K. J., Williams, S., Howard, R., Frackowiak, R. S., & Turner, R. (1996). Movement-related effects in fMRI time-series. *Magn Reson Med, 35*(3), 346-355.

Friston, K. J. A., J.T. ; Kiebel, S.J. ; Nichols, T.E. ; Penny, W.D. (2007). *Statistical Parametric Mapping The Analysis of Functional Brain Imaging*. Hungary: Elsevier.

Fuster, J. M. (1998). Linkage at the top. *Neuron, 21*(6), 1223-1224.

Fuster, J. M., & Oxford University, P. (2003). *Cortex and mind : unifying cognition*. Oxford ; New York: Oxford University Press.

Gabrieli, J. D., Cohen, N. J., & Corkin, S. (1988). The impaired learning of semantic knowledge following bilateral medial temporal-lobe resection. *Brain Cogn, 7*(2), 157-177.

Gadian, D. G. (1995). *NMR and its Applications to Living Systems*. New York: Oxford University Press Inc.

Gaillard, W. D., Balsamo, L., Xu, B., McKinney, C., Papero, P. H., Weinstein, S., . . . Theodore, W. H. (2004). fMRI language task panel improves determination of language dominance. *Neurology, 63*(8), 1403-1408.

Gaillard, W. D., Balsamo, L. M., Ibrahim, Z., Sachs, B. C., & Xu, B. (2003). fMRI identifies regional specialization of neural networks for reading in young children. *Neurology, 60*(1), 94-100.

Gaillard, W. D., Berl, M. M., Moore, E. N., Ritzl, E. K., Rosenberger, L. R., Weinstein, S. L., . . . Theodore, W. H. (2007). Atypical language in lesional and nonlesional complex partial epilepsy. *Neurology, 69*(18), 1761-1771.

Gaillard, W. D., Grandin, C. B., & Xu, B. (2001). Developmental aspects of pediatric fMRI: considerations for image acquisition, analysis, and interpretation. *NeuroImage, 13*(2), 239-249.

Gaillard, W. D., Hertz-Pannier, L., Mott, S. H., Barnett, A. S., LeBihan, D., & Theodore, W. H. (2000). Functional anatomy of cognitive development: fMRI of verbal fluency in children and adults. *Neurology, 54*(1), 180-185.

Gaillard, W. D., Pugliese, M., Grandin, C. B., Braniecki, S. H., Kondapaneni, P., Hunter, K., . . . Basso, G. (2001). Cortical localization of reading in normal children: an fMRI language study. *Neurology, 57*(1), 47-54.

Gardner, M. F. (1990). *EO-WPVT-R.: Expressive One-Word Picture Vocabulary Test, Revised. Record Forms*: Academic Therapy Publications.

Gardner, W. J. (1941). Injection of procaine into the brain to locate speech area in left-handed persons. *Archives of Neurology and Psychiatry, 46*(6), 1035-1038.

Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of Statistical Maps in Functional Neuroimaging Using the False Discovery Rate. *NeuroImage, 15*(4), 870-878.

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation, 12*(4), 395-427.

Gioia, C., Isquith, P., Guy, S. C., & Kenworthy, L. (2000). *Behavior Rating Inventory of Executive Function*. Odessa, FL: Psychological Assessment Resources.

Giussani, C., Roux, F. E., Ojemann, J., Sganzerla, E. P., Pirillo, D., & Papagno, C. (2010). Is preoperative functional magnetic resonance imaging reliable for language areas mapping in brain tumor surgery? Review of language functional magnetic resonance imaging and direct cortical stimulation correlation studies. *Neurosurgery, 66*(1), 113-120.

Gleissner, U., Sassen, R., Schramm, J., Elger, C. E., & Helmstaedter, C. (2005). Greater functional recovery after temporal lobe epilepsy surgery in children. *Brain, 128*(Pt 12), 2822-2829.

Gracco, V. L., Tremblay, P., & Pike, B. (2005). Imaging speech production using fMRI. *NeuroImage, 26*(1), 294-301.

Hagoort, P. (2013). MUC (Memory, Unification, Control) and beyond. *Front Psychol, 4*, 416.

Haller, S., Radue, E. W., Erb, M., Grodd, W., & Kircher, T. (2005). Overt sentence production in event-related fMRI. *Neuropsychologia, 43*(5), 807-814.

Hallowell, L. M., Stewart, S. E., de Amorim, E. S. C. T., & Ditchfield, M. R. (2008). Reviewing the process of preparing children for MRI. *Pediatr Radiol, 38*(3), 271-279.

Hamberger, M. J., Goodman, R. R., Perrine, K., & Tamny, T. (2001). Anatomic dissociation of auditory and visual naming in the lateral temporal cortex. *Neurology, 56*(1), 56-61.

Hamberger, M. J., Habeck, C. G., Pantazatos, S. P., Williams, A. C., & Hirsch, J. (2014). Shared space, separate processes: Neural activation patterns for auditory description and visual object naming in healthy adults. *Human Brain Mapping, 35*(6), 2507-2520.

Hamberger, M. J., McClelland, S., 3rd, McKhann, G. M., 2nd, Williams, A. C., & Goodman, R. R. (2007). Distribution of auditory and visual naming sites in nonlesional temporal lobe epilepsy patients and patients with space-occupying temporal lobe lesions. *Epilepsia, 48*(3), 531-538.

Hamberger, M. J., Seidel, W. T., Goodman, R. R., Perrine, K., & McKhann, G. M. (2003). Temporal lobe stimulation reveals anatomic distinction between auditory naming processes. *Neurology, 60*(9), 1478-1483.

Hamberger, M. J., Seidel, W. T., Goodman, R. R., Williams, A., Perrine, K., Devinsky, O., & McKhann, G. M., 2nd. (2007). Evidence for cortical reorganization of language in patients with hippocampal sclerosis. *Brain, 130*(Pt 11), 2942-2950.

Harris, J. J., Reynell, C., & Attwell, D. (2011). The physiology of developmental chnages in BOLD functional imaging signals. *Dev Cogn Neurosci, 1*, 199-216.

Harvey, A. S., Cross, J. H., Shinnar, S., Mathern, G. W., & Taskforce, I. P. E. S. S. (2008). Defining the spectrum of international practice in pediatric epilepsy surgery patients. *Epilepsia, 49*(1), 146-155.

Heim, S., Eickhoff, S. B., & Amunts, K. (2009). Different roles of cytoarchitectonic BA 44 and BA 45 in phonological and semantic verbal fluency as revealed by dynamic causal modelling. *NeuroImage, 48*(3), 616-624.

Heim, S., Eickhoff, S. B., Friederici, A. D., & Amunts, K. (2009). Left cytoarchitectonic area 44 supports selection in the mental lexicon during language production. *Brain Struct Funct, 213*(4-5), 441-456.

Henson, R., Rugg, M. D., & Friston, K. J. (2001). The choice of basis functions in event-related fMRI. *NeuroImage, 13*(6), S149-S149.

Hermann, B., Jones, J., Dabbs, K., Allen, C. A., Sheth, R., Fine, J., . . . Seidenberg, M. (2007). The frequency, complications and aetiology of ADHD in new onset paediatric epilepsy. *Brain, 130*(Pt 12), 3135-3148.

Hickok, G. (2009). The functional neuroanatomy of language. *Phys Life Rev, 6*(3), 121-143.

Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends Cogn Sci, 4*(4), 131-138.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition, 92*(1-2), 67-99.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nat Rev Neurosci, 8*(5), 393-402.

Hillis, A. E., & Caramazza, A. (1991). Category-Specific Naming and Comprehension Impairment - a Double Dissociation. *Brain, 114*, 2081-2094.

Hirsch, E., & Arzimanoglou, A. (2004). [Children with drug-resistant partial epilepsy: criteria for the identification of surgical candidates]. *Rev Neurol (Paris), 160 Spec No 1*, 5S210-219.

Hodges, J. R., & Patterson, K. (2007). Semantic dementia: a unique clinicopathological syndrome. *Lancet Neurol, 6*(11), 1004-1014.

Hodges, J. R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Semantic Dementia - Progressive Fluent Aphasia with Temporal-Lobe Atrophy. *Brain, 115*, 1783-1806.

Holland, R., & Lambon Ralph, M. A. (2010). The anterior temporal lobe semantic hub is a part of the language neural network: selective disruption of irregular past tense verbs by rTMS. *Cereb Cortex, 20*(12), 2771-2775.

Holland, S. K., Plante, E., Weber Byars, A., Strawsburg, R. H., Schmithorst, V. J., & Ball, W. S., Jr. (2001). Normal fMRI brain activation patterns in children performing a verb generation task. *NeuroImage, 14*(4), 837-843.

Holland, S. K., Vannest, J., Mecoli, M., Jacola, L. M., Tillema, J. M., Karunanayaka, P. R., . . . Byars, A. W. (2007). Functional MRI of language lateralization during development in children. *Int J Audiol, 46*(9), 533-551.

Holmes, S. J., & Ellis, A. W. (2007). Age of acquisition and typicality effects in three object processing tasks. *Visual Cognition, 13*(7-8).

Homae, F., Hashimoto, R., Nakajima, K., Miyashita, Y., & Sakai, K. L. (2002). From perception to sentence comprehension: the convergence of auditory and visual information of language in the left inferior frontal cortex. *NeuroImage, 16*(4), 883-900.

Hornak, J. P. (1996). The Basics of MRI. from http://www.cis.rit.edu/htbooks/mri/index.html

Huttenlocher, J., & Kubicek, L. F. (1983). The source of relatedness effects on naming latency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*(3), 486.

Hutton, C., Bork, A., Josephs, O., Deichmann, R., Ashburner, J., & Turner, R. (2002). Image distortion correction in fMRI: A quantitative evaluation. *NeuroImage, 16*(1), 217-240.

Indefrey, P., Brown, C. M., Hellwig, F., Amunts, K., Herzog, H., Seitz, R. J., & Hagoort, P. (2001). A neural correlate of syntactic encoding during speech production. *Proceedings of the National Academy of Sciences, 98*(10), 5933-5936.

Jacobs, J. R., Hawco, C., Kobayashi, E., Boor, R., LeVan, P., Stephni, U., . . . Gotman, J. (2008). Variability of the haemodynamic response function as a functin of age and frequency of epileptic discharge in children with epilepsy. *NeuroImage, 40*, 601-614.

Jacquemot, C., & Scott, S. K. (2006). What is the relationship between phonological short-term memory and speech processing? *Trends Cogn Sci, 10*(11), 480-486.

Janecek, J. K., Swanson, S. J., Sabsevitz, D. S., Hammeke, T. A., Raghavan, M., M, E. R., & Binder, J. R. (2013). Language lateralization by fMRI and Wada testing in 229 patients with epilepsy: rates and predictors of discordance. *Epilepsia, 54*(2), 314-322.

Jansen, F. E., Jennekens-Schinkel, A., Van Huffelen, A. C., Van Veelen, W. M., Van Rijen, C. P., Alpherts, W. C., . . . Van Nieuwenhuizen, O. (2002). Diagnostic significance of Wada procedure in very young children and children with developmental delay. *Eur J Paediatr Neurol, 6*(6), 315-320.

JEC. (2010). Joint Epilepsy Council: National incidence and prevelance of epilepsy.

Jefferies, E., Patterson, K., & Ralph, M. A. (2008). Deficits of knowledge versus executive control in semantic cognition: insights from cued naming. *Neuropsychologia, 46*(2), 649-658.

Johnson, M. H. (2001). Functional brain development in humans. *Nat Rev Neurosci, 2*(7), 475-483.

Johnson, M. H. (2011). Interactive specialization: a domain-general framework for human functional brain development? *Dev Cogn Neurosci, 1*(1), 7-21.

Johnston, R. A., Dent, K., Humphreys, G. W., & Barry, C. (2010). British-English norms and naming times for a set of 539 pictures: The role of age of acquisition. *Behav Res Methods, 42*(2), 461-469.

Josephs, O., Turner, R., & Friston, K. (1997). Event-related fMRI. *Human Brain Mapping, 5*(4), 243-248.

Just, M. A., Newman, S. D., Keller, T. A., McEleney, A., & Carpenter, P. A. (2004). Imagery in sentence comprehension: an fMRI study. *NeuroImage, 21*(1), 112-124.

Kang, H. C., Burgund, E. D., Lugar, H. M., Petersen, S. E., & Schlaggar, B. L. (2003). Comparison of functional activation foci in children and adults using a common stereotactic space. *NeuroImage, 19*(1), 16-28.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci, 17*(11), 4302-4311.

Karunanayaka, P., Schmithorst, V. J., Vannest, J., Szaflarski, J. P., Plante, E., & Holland, S. K. (2010). A group independent component analysis of covert verb generation in children: a functional magnetic resonance imaging study. *NeuroImage, 51*(1), 472-487.

Kim, K. K., Privitera, M. D., & Szaflarski, J. P. (2011). Lessons learned from a comparison of language localisation using fMRI and electrocortical mapping: case studies of neocortical epilepsy patients. *Epileptic Disord, 13*(4), 368-374.

Kircher, T. T. J., Brammer, M. J., Williams, S. C. R., & McGuire, P. K. (2000). Lexical retrieval during fluent speech production: an fMRI study. *Neuroreport, 11*(18), 4093-4096.

Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Increased brain activity in frontal and parietal cortex underlies the development of visuospatial working memory capacity during childhood. *J Cogn Neurosci, 14*(1), 1-10.

Korkman, M., Granstrom, M. L., Kantola-Sorsa, E., Gaily, E., Paetau, R., Liukkonen, E., . . . Blomstedt, G. (2005). Two-year follow-up of intelligence after pediatric epilepsy surgery. *Pediatr Neurol, 33*(3), 173-178.

Krauss, G. L., Fisher, R., Plate, C., Hart, J., Uematsu, S., Gordon, B., & Lesser, R. P. (1996). Cognitive effects of resecting basal temporal language areas. *Epilepsia, 37*(5), 476-483.

Kummerer, D., Hartwigsen, G., Kellmeyer, P., Glauche, V., Mader, I., Kloppel, S., . . . Saur, D. (2013). Damage to ventral and dorsal language pathways in acute aphasia. *Brain, 136*(Pt 2), 619-629.

Kurczek, J., Brown-Schmidt, S., & Duff, M. (2013). Hippocampal Contributions to Language: Evidence of Referential Processing Deficits in Amnesia. *Journal of Experimental Psychology-General, 142*(4), 1346-1354.

Kwan, P., Arzimanoglou, A., Berg, A. T., Brodie, M. J., Allen Hauser, W., Mathern, G., . . . French, J. (2010). Definition of drug resistant epilepsy: consensus proposal by the ad hoc Task Force of the ILAE Commission on Therapeutic Strategies. *Epilepsia, 51*(6), 1069-1077.

Kwong, K. K., Belliveau, J. W., Chesler, D. A., Goldberg, I. E., Weisskoff, R. M., Poncelet, B. P., . . . et al. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc Natl Acad Sci U S A, 89*(12), 5675-5679.

Lachman, R. (1973). Uncertainty effects on time to access the internal lexicon. *Journal of Experimental Psychology, 99*(2), 199.

Lachman, R., Shaffer, J. P., & Hennrikus, D. (1974). Language and cognition: Effects of stimulus codability, name-word frequency, and age of acquisition on lexical reaction time 1. *13*(6), 613–625.

Lambon Ralph, M. A., Pobric, G., & Jefferies, E. (2009). Conceptual knowledge is underpinned by the temporal pole bilaterally: convergent evidence from rTMS. *Cereb Cortex, 19*(4), 832-838.

Lebel, C., Walker, L., Leemans, A., Phillips, L., & Beaulieu, C. (2008). Microstructural maturation of the human brain from childhood to adulthood. *NeuroImage, 40*(3), 1044-1055.

Lee, B. C., Kuppusamy, K., Grueneich, R., El-Ghazzawy, O., Gordon, R. E., Lin, W., & Haacke, E. M. (1999). Hemispheric language dominance in children demonstrated by functional magnetic resonance imaging. *J Child Neurol, 14*(2), 78-82.

Lemieux, L., Salek-Haddadi, A., Lund, T. E., Laufs, H., & Carmichael, D. (2007). Modelling large motion events in fMRI studies of patients with epilepsy. *Magn Reson Imaging, 25*(6), 894-901.

Leniger-Follert, E., & Hossmann, K. A. (1979). Simultaneous measurements of microflow and evoked potentials in the somatomotor cortex of the cat brain during specific sensory activation. *Pflugers Arch, 380*(1), 85-89.

Lenroot, R. K., & Giedd, J. N. (2006). Brain development in children and adolescents: insights from anatomical magnetic resonance imaging. *Neuroscience and Biobehavioural Reviews, 30*(6), 718-729.

Lettori, D., Battaglia, D., Sacco, A., Veredice, C., Chieffo, D., Massimi, L., . . . Guzzetta, F. (2008). Early hemispherectomy in catastrophic epilepsy: a neuro-cognitive and epileptic long-term follow-up. *Seizure, 17*(1), 49-63.

Leube, D. T., Erb, M., Grodd, W., Bartels, M., & Kircher, T. T. J. (2001). Activation of right fronto-temporal cortex characterizes the 'living' category in semantic processing. *Cognitive Brain Research, 12*(3), 425-430.

Levelt, W. J. (1999). Producing spoken language: A blueprint of the speaker *The neurocognition of language* (pp. 83-122).

Lidzba, K., Schwilling, E., Grodd, W., Krageloh-Mann, I., & Wilke, M. (2011). Language comprehension vs. language production: age effects on fMRI activation. *Brain and Language, 119*(1), 6-15.

Lieberman, M. D., & Cunningham, W. A. (2009). Type I and Type II error concerns in fMRI research: re-balancing the scale. *Soc Cogn Affect Neurosci, 4*(4), 423-428.

Liegeois, F., Connelly, A., Cross, J. H., Boyd, S. G., Gadian, D. G., Vargha-Khadem, F., & Baldeweg, T. (2004). Language reorganization in children with early-onset lesions of the left hemisphere: an fMRI study. *Brain, 127*(Pt 6), 1229-1236.

Liegeois, F., Connelly, A., Salmond, C. H., Gadian, D. G., Vargha-Khadem, F., & Baldeweg, T. (2002). A Direct Test for Lateralization of Language Activation using fMRI: Comparison with Invasive Assessments in Children with Epilepsy. *NeuroImage, 17,* 1861-1867.

Loddenkemper, T., Holland, K. D., Stanford, L. D., Kotagal, P., Bingaman, W., & Wyllie, E. (2007). Developmental outcome after epilepsy surgery in infancy. *Pediatrics, 119*(5), 930-935.

Logothetis, N. K. (2003). The Underpinnings of the BOLD Functional Magnetic Resonance Imaging Signal. *The Journal of Neuroscience*.

Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature, 453*(7197), 869-878.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature, 412*(6843), 150-157.

Luders, H., Hahn, J., Lesser, R. P., Dinner, D. S., Morris, H. H., 3rd, Wyllie, E., . . . Skipper, G. (1989). Basal temporal subdural electrodes in the evaluation of patients with intractable epilepsy. *Epilepsia, 30*(2), 131-142.

Luders, H., Lesser, R. P., Hahn, J., Dinner, D. S., Morris, H., Resor, S., & Harrison, M. (1986). Basal temporal language area demonstrated by electrical stimulation. *Neurology, 36*(4), 505-510.

Luders, H., Lesser, R. P., Hahn, J., Dinner, D. S., Morris, H. H., Wyllie, E., & Godoy, J. (1991). Basal temporal language area. *Brain, 114 (Pt 2)*, 743-754.

Lurito, J. T., Lowe, M. J., Sartorius, C., & Mathews, V. P. (2000). Comparison of fMRI and intraoperative direct cortical stimulation in localization of receptive language areas. *J Comput Assist Tomogr, 24*(1), 99-105.

Majerus, S. (2013). Language repetition and short-term memory: an integrative framework. *Front Hum Neurosci, 7*, 357.

Makris, N., & Pandya, D. N. (2009). The extreme capsule in humans and rethinking of the language circuitry. *Brain Struct Funct, 213*(3), 343-358.

Makris, N., Papadimitriou, G. M., Kaiser, J. R., Sorg, S., Kennedy, D. N., & Pandya, D. N. (2009). Delineation of the middle longitudinal fascicle in humans: a quantitative, in vivo, DT-MRI study. *Cerebral Cortex, 19*(4), 777-785.

Makris, N., Papadimitriou, G. M., Sorg, S., Kennedy, D. N., Caviness, V. S., & Pandya, D. N. (2007). The occipitofrontal fascicle in humans: a quantitative, in vivo, DT-MRI study. *NeuroImage, 37*(4), 1100-1111.

Makuuchi, M., Bahlmann, J., Anwander, A., & Friederici, A. D. (2009). Segregating the core computational faculty of human language from working memory. *Proc Natl Acad Sci U S A, 106*(20), 8362-8367.

Makuuchi, M., & Friederici, A. D. (2013). Hierarchical functional connectivity between the core language system and the working memory system. *Cortex, 49*(9), 2416-2423.

Maldjian, J. A., Laurienti, P. J., & Burdette, J. H. (2004). Precentral gyrus discrepancy in electronic versions of the Talairach atlas. *NeuroImage, 21*(1), 450-455.

Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage, 19*(3), 1233-1239.

Maldonado, I. L., Moritz-Gasser, S., & Duffau, H. (2011). Does the left superior longitudinal fascicle subserve language semantics? A brain electrostimulation study. *Brain Struct Funct, 216*(3), 263-274.

Malonek, D., & Grinvald, A. (1996). Interactions between electrical activity and cortical microcirculation revealed by imaging spectroscopy: implications for functional brain mapping. *Science, 272*(5261), 551-554.

Marteau, T. M., & Bekker, H. (1992). The Development of a 6-Item Short-Form of the State Scale of the Spielberger State Trait Anxiety Inventory (Stai). *British Journal of Clinical Psychology, 31*, 301-306.

McCarthy, G., Puce, A., Gore, J. C., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *J Cogn Neurosci, 9*(5), 605-610.

Mechelli, A., Henson, R. N., Price, C. J., & Friston, K. J. (2003). Comparing event-related and epoch analysis in blocked design fMRI. *NeuroImage, 18*(3), 806-810.

Mesulam, M. M. (1990). Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Ann Neurol, 28*(5), 597-613.

Mion, M., Patterson, K., Acosta-Cabronero, J., Pengas, G., Izquierdo-Garcia, D., Hong, Y. T., . . . Nestor, P. J. (2010). What the left and right anterior fusiform gyri tell us about semantic memory. *Brain, 133*(11), 3256-3268.

Muller, R. A., Rothermel, R. D., Behen, M. E., Muzik, O., Mangner, T. J., & Chugani, H. T. (1997). Receptive and expressive language activations for sentences: A PET study. *Neuroreport, 8*(17), 3767-3770.

Mummery, C. J., Patterson, K., Price, C. J., Ashburner, J., Frackowiak, R. S., & Hodges, J. R. (2000). A voxel-based morphometry study of semantic dementia: relationship between temporal lobe atrophy and semantic memory. *Ann Neurol, 47*(1), 36-45.

Noppeney, U., & Price, C. J. (2002). A PET study of stimulus- and task-induced semantic processing. *NeuroImage, 15*(4), 927-935.

Northam, G. B., Liegeois, F., Tournier, J. D., Croft, L. J., Johns, P. N., Chong, W. K., . . . Baldeweg, T. (2012). Interhemispheric temporal lobe connectivity predicts language impairment in. *Brain, 135*(Pt 12), 3781-3798.

Nunez, S. C., Dapretto, M., Katzir, T., Starr, A., Bramen, J., Kan, E., . . . Sowell, E. R. (2011). fMRI of syntactic processing in typically developing children: structural correlates in the inferior frontal gyrus. *Dev Cogn Neurosci, 1*(3), 313-323.

Obleser, J., Zimmermann, J., Van Meter, J., & Rauschecker, J. P. (2007). Multiple stages of auditory speech perception reflected in event-related FMRI. *Cereb Cortex, 17*(10), 2251-2257.

Ogawa, S., & Lee, T. M. (1990). Magnetic-Resonance-Imaging of Blood-Vessels at High Fields - Invivo and Invitro Measurements and Image Simulation. *Magnetic Resonance in Medicine, 16*(1), 9-18.

Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain Magnetic-Resonance-Imaging with Contrast Dependent on Blood Oxygenation. *Proc Natl Acad Sci U S A, 87*(24), 9868-9872.

Ogawa, S., Lee, T. M., Nayak, A. S., & Glynn, P. (1990). Oxygenation-Sensitive Contrast in Magnetic-Resonance Image of Rodent Brain at High Magnetic-Fields. *Magnetic Resonance in Medicine, 14*(1), 68-78.

Ogawa, S., Menon, R. S., Tank, D. W., Kim, S. G., Merkle, H., Ellermann, J. M., & Ugurbil, K. (1993). Functional brain mapping by blood oxygenation level-dependent contrast magnetic resonance imaging. A comparison of signal characteristics with a biophysical model. *Biophys J, 64*(3), 803-812.

Ogawa, S., Tank, D. W., Menon, R., Ellermann, J. M., Kim, S. G., Merkle, H., & Ugurbil, K. (1992). Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. *Proc Natl Acad Sci U S A, 89*(13), 5951-5955.

Ojemann, J. G., Ojemann, G. A., & Lettich, E. (2002). Cortical stimulation mapping of language cortex by using a verb generation task: effects of learning and comparison to mapping based on object naming. *J Neurosurg, 97*(1), 33-38.

Okada, T., Tanaka, S., Nakai, T., Nishizawa, S., Inui, T., Sadato, N., . . . Konishi, J. (2000). Naming of animals and tools: a functional magnetic resonance imaging study of categorical differences in the human brain areas commonly used for naming visually presented objects. *Neurosci Lett, 296*(1), 33-36.

Oldfield, R. C. (1971). The Assessment and Analysis of Handedness: The Edinburgh Inventory. *Neuropsychologia, 9*(1), 97-113.

Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Q J Exp Psychol, 17*(4), 273-281.

Opitz, B., & Friederici, A. D. (2004). Brain correlates of language learning: The neuronal dissociation of rule-based versus similarity-based learning. *Journal of Neuroscience, 24*(39), 8436-8440.

Orton, S. T. (1935). *Reading, Writing, and Speech Problems in Children and Selected Papers*. Baltimore: The International Dyslexia Association.

Pahs, G., Rankin, P., Helen Cross, J., Croft, L., Northam, G. B., Liegeois, F., . . . Baldeweg, T. (2013). Asymmetry of planum temporale constrains interhemispheric language plasticity in children with focal epilepsy. *Brain, 136*(Pt 10), 3163-3175.

Parkinson, G. M. (2002). High incidence of language disorder in children with focal epilepsies. *Developmental Medicine & Child Neurology, 44*(8), 50-56.

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat Rev Neurosci, 8*(12), 976-987.

Penfield, W., & Boldrey, E. (1937). Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. *Brain, 60*, 389-443.

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature, 331*(6157), 585-589.

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1989). Positron emission tomographic studies of the processing of singe words. *J Cogn Neurosci, 1*(2), 153-170.

Picot, M. C., Baldy-Moulinier, M., Daures, J. P., Dujols, P., & Crespel, A. (2008). The prevalence of epilepsy and pharmacoresistant epilepsy in adults: a population-based study in a Western European country. *Epilepsia, 49*(7), 1230-1238.

Plante, E., Holland, S. K., & Schmithorst, V. J. (2006). Prosodic processing by children: an fMRI study. *Brain Lang, 97*(3), 332-342.

Plante, E., Schmithorst, V. J., Holland, S. K., & Byars, A. W. (2006). Sex differences in the activation of language cortex during childhood. *Neuropsychologia, 44*(7), 1210-1221.

Poeppel, D., Emmorey, K., Hickok, G., & Pylkkanen, L. (2012). Towards a new neurobiology of language. *J Neurosci, 32*(41), 14125-14131.

Poeppel, D., & Hickok, G. (2004). Towards a new functional anatomy of language. *Cognition, 92*(1-2), 1-12.

Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *NeuroImage, 10*(1), 15-35.

Potvin, D., & Nussbaum, N. (2013). Methods of lateralization of language in pre-surgical assessments of children with intractable epilepsy. *Journal of Pediatric Epilepsy, 2*, 215-221.

Pravata, E., Sestieri, C., Mantini, D., Briganti, C., Colicchio, G., Marra, C., . . . Caulo, M. (2011). Functional connectivity MR imaging of the language network in patients with drug-resistant epilepsy. *American Journal of Neuroradiology, 32*(3), 532-540.

Pressdee, D., May, L., Eastman, E., & Grier, D. (1997). The use of play therapy in the preparation of children undergoing MR imaging. *Clinical Radiology, 52*(12), 945-947.

Price, C. J. (2010). The anatomy of language: a review of 100 fMRI studies published in 2009. *Ann N Y Acad Sci, 1191*, 62-88.

Price, C. J. (2012). A review and synthesis of the first 20years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage, 62*(2), 816-847.

Price, C. J., & Devlin, J. T. (2003). The myth of the visual word form area. *NeuroImage, 19*(3), 473-481.

Price, C. J., & Devlin, J. T. (2011). The interactive account of ventral occipitotemporal contributions to reading. *Trends Cogn Sci, 15*(6), 246-253.

Price, C. J., & MacSweeney, M. (2011). A generative model of speech production in Broca's and Wernicke's areas. *Frontiers in Language Sciences, 2*.

Price, C. J., Veltman, D. J., Ashburner, J., Josephs, O., & Friston, K. J. (1999). The critical relationship between the timing of stimulus presentation and data acquisition in blocked designs with fMRI. *NeuroImage, 10*(1), 36-44.

Protzner, A. B., & McAndrews, M. P. (2011). Network alterations supporting word retrieval in patients with medial temporal lobe epilepsy. *J Cogn Neurosci, 23*(9), 2605-2619.

Pulsifer, M. B., Brandt, J., Salorio, C. F., Vining, E. P., Carson, B. S., & Freeman, J. M. (2004). The cognitive outcome of hemispherectomy in 71 children. *Epilepsia, 45*(3), 243-254.

Pulvermüller, F., & Fadiga, L. (2010). Active perception: sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience, 11*(5), 351-360.

Raettig, T., & Kotz, S. A. (2008). Auditory processing of different types of pseudo-words: an event-related fMRI study. *NeuroImage, 39*(3), 1420-1428.

Raschle, N. M., Lee, M., Buechler, R., Christodoulou, J. A., Chang, M., Vakil, M., . . . Gaab, N. (2009). Making MR imaging child's play - pediatric neuroimaging protocol, guidelines and procedure. *J Vis Exp*(29).

Richardson, F. M., Seghier, M. L., Leff, A. P., Thomas, M. S., & Price, C. J. (2011). Multiple routes from occipital to temporal cortices during reading. *J Neurosci, 31*(22), 8239-8247.

Rolheiser, T., Stamatakis, E. A., & Tyler, L. K. (2011). Dynamic processing in the human language system: synergy between the arcuate fascicle and extreme capsule. *Journal of Neuroscience, 31*(47), 16949-16957.

Rosenberger, L. R., Zeck, J., Berl, M. M., Moore, E. N., Ritzl, E. K., Shamim, S., . . . Gaillard, W. D. (2009). Interhemispheric and intrahemispheric language reorganization in complex partial epilepsy. *Neurology, 72*(21), 1830-1836.

Roux, F. E., Boulanouar, K., Lotterie, J. A., Mejdoubi, M., LeSage, J. P., & Berry, I. (2003). Language functional magnetic resonance imaging in preoperative assessment of language areas: correlation with direct cortical stimulation. *Neurosurgery, 52*(6), 1335-1345; discussion 1345-1337.

Ruge, M. I., Victor, J., Hosain, S., Correa, D. D., Relkin, N. R., Tabar, V., . . . Hirsch, J. (1999). Concordance between functional magnetic resonance imaging and intraoperative language mapping. *Stereotact Funct Neurosurg, 72*(2-4), 95-102.

Rushworth, M. F., Walton, M. E., Kennerley, S. W., & Bannerman, D. M. (2004). Action sets and decisions in the medial frontal cortex. *Trends Cogn Sci, 8*(9), 410-417.

Rutten, G. J., Ramsey, N. F., van Rijen, P. C., Noordmans, H. J., & van Veelen, C. W. (2002). Development of a functional magnetic resonance imaging protocol for intraoperative localization of critical temporoparietal language areas. *Ann Neurol, 51*(3), 350-360.

Rutter, M., Graham, P. J., & Yule, W. (1970). *A neuropsychiatric study in childhood*. London: Heinemann Medical for Spastics International Medical Publications.

Sabsevitz, D. S., Swanson, S. J., Hammeke, T. A., Spanaki, M. V., Possing, E. T., Morris, G. L., 3rd, . . . Binder, J. R. (2003). Use of preoperative functional neuroimaging to predict language deficits from epilepsy surgery. *Neurology, 60*(11), 1788-1792.

Saur, D., Kreher, B. W., Schnell, S., Kummerer, D., Kellmeyer, P., Vry, M. S., . . . Weiller, C. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Science USA, 105*(46), 18035-18040.

Saur, D., Schelter, B., Schnell, S., Kratochvil, D., Kupper, H., Kellmeyer, P., . . . Weiller, C. (2010). Combining functional and anatomical connectivity reveals brain networks for auditory language comprehension. *NeuroImage, 49*(4), 3187-3197.

Schaffler, L., Luders, H. O., & Beck, G. J. (1996). Quantitative comparison of language deficits produced by extraoperative electrical stimulation of Broca's, Wernicke's, and basal temporal language areas. *Epilepsia, 37*(5), 463-475.

Schlund, M. W., Cataldo, M. F., Siegle, G. J., Ladouceur, C. D., Silk, J. S., Forbes, E. E., . . . Ryan, N. D. (2011). Pediatric functional magnetic resonance neuroimaging: tactics for encouraging task compliance. *Behav Brain Funct, 7*, 10.

Schmahmann, J. D., & Pandya, D. N. (2007a). Cerebral white matter--historical evolution of facts and notions concerning the organization of the fiber pathways of the brain. *J Hist Neurosci, 16*(3), 237-267.

Schmahmann, J. D., & Pandya, D. N. (2007b). The complex history of the fronto-occipital fasciculus. *J Hist Neurosci, 16*(4), 362-377.

Schmahmann, J. D., Pandya, D. N., Wang, R., Dai, G., D'Arceuil, H. E., de Crespigny, A. J., & Wedeen, V. J. (2007). Association fibre pathways of the brain: parallel observations from diffusion spectrum imaging and autoradiography. *Brain, 130*(Pt 3), 630-653.

Schmithorst, V. J., Holland, S. K., & Plante, E. (2006). Cognitive modules utilized for narrative comprehension in children: a functional magnetic resonance imaging study. *NeuroImage, 29*(1), 254-266.

Schmithorst, V. J., Holland, S. K., & Plante, E. (2007). Object identification and lexical/semantic access in children: a functional magnetic resonance imaging study of word-picture matching. *Human Brain Mapping, 28*(10), 1060-1074.

Schwartz, T. H. (2009). Cortical dyplasia: complete resection correlates with outcome ... But, complete resection of what? *Epilepsy Curr, 9*(4), 100-102.

Schwartz, T. H., Devinsky, O., Doyle, W., & Perrine, K. (1998). Preoperative predictors of anterior temporal language areas. *J Neurosurg, 89*(6), 962-970.

Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain, 123* (Pt 12), 2400-2406.

Scott, S. K., Rosen, S., Lang, H., & Wise, R. J. (2006). Neural correlates of intelligibility in speech investigated with noise vocoded speech--a positron emission tomography study. *J Acoust Soc Am, 120*(2), 1075-1083.

Seghier, M. L. (2008). Laterality index in functional MRI: methodological issues. *Magn Reson Imaging, 26*(5), 594-601.

Seghier, M. L., Lee, H. L., Schofield, T., Ellis, C. L., & Price, C. J. (2008). Inter-subject variability in the use of two different neuronal networks for reading aloud familiar words. *NeuroImage, 42*(3), 1226-1236.

Seghier, M. L., & Price, C. J. (2011). Explaining left lateralization for words in the ventral occipitotemporal cortex. *J Neurosci, 31*(41), 14745-14753.

Seghier, M. L., & Price, C. J. (2013). Dissociating frontal regions that co-lateralize with different ventral occipitotemporal regions during word processing. *Brain Lang, 126*(2), 133-140.

Semel, E., Wiig, E. H., & Secord, W. (2003a). *Clinical Evaluation of Language Fundamentals-IV*: Marickville: Harcourt Assessment.

Semel, E., Wiig, E. H., & Secord, W. (2003b). *Clinical evaluation of language fundamentals, fourth edition (CELF-4)*. Toronto, Canada: The Psychological Corporation/A Harcourt Assessment Company.

Shaywitz, B. A., Skudlarski, P., Holahan, J. M., Marchione, K. E., Constable, R. T., Fulbright, R. K., . . . Shaywitz, S. E. (2007). Age-related changes in reading systems of dyslexic children. *Ann Neurol, 61*(4), 363-370.

Sherman, E. M., Wiebe, S., Fay-McClymont, T. B., Tellez-Zenteno, J., Metcalfe, A., Hernandez-Ronquillo, L., . . . Jette, N. (2011). Neuropsychological outcomes after epilepsy surgery: systematic review and pooled estimates. *Epilepsia, 52*(5), 857-869.

Sherman, E. M. S., Alberta Children's Hospital, C., Alberta, Canada, University of British Columbia, V., British Columbia, Canada, Slick, D. J., Alberta Children's Hospital, C., Alberta, Canada, Connolly, M. B., . . . BC Children's Hospital, V., British Columbia, Canada. (2007). ADHD, Neurological Correlates and Health‐related Quality of Life in Severe Pediatric Epilepsy. *Epilepsia, 48*(6), 1083-1091.

Skirrow, C., Cross, J. H., Cormack, F., Harkness, W., Vargha-Khadem, F., & Baldeweg, T. (2011). Long-term intellectual outcome after temporal lobe surgery in childhood. *Neurology, 76*(15), 1330-1337.

Sladky, R., Friston, K. J., Trostl, J., Cunnington, R., Moser, E., & Windischberger, C. (2011). Slice-timing effects and their correction in functional MRI. *NeuroImage, 58*(2), 588-594.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of experimental psychology Human learning and memory, 6*(2), 174-215.

Snodgrass, J. G., & Yuditsky, T. (1996). Naming times for the Snodgrass and Vanderwart pictures. *Behavior Research Methods, Instruments, & Computers, 28*(4), 516-536.

Soros, P., Sokoloff, L. G., Bose, A., McIntosh, A. R., Graham, S. J., & Stuss, D. T. (2006). Clustered functional MRI of overt speech production. *NeuroImage, 32*(1), 376-387.

Spencer, S., & Huh, L. (2008). Outcomes of epilepsy surgery in adults and children. *Lancet Neurol, 7*(6), 525-537.

Suarez, R. O., Whalen, S., Nelson, A. P., Tie, Y., Meadows, M. E., Radmanesh, A., & Golby, A. J. (2009). Threshold-independent functional MRI determination of language dominance: a validation study against clinical gold standards. *Epilepsy & Behavior: E&B, 16*(2), 288-297.

Sundermann, B., & Pfleiderer, B. (2012). Functional connectivity profile of the human inferior frontal junction: involvement in a cognitive control network. *BMC Neurosci, 13*, 119.

Szabo, C. A., Wyllie, E., Stanford, L. D., Geckler, C., Kotagal, P., Comair, Y. G., & Thornton, A. E. (1998). Neuropsychological effect of temporal lobe resection in preadolescent children with epilepsy. *Epilepsia, 39*(8), 814-819.

Szaflarski, J. P., Altaye, M., Rajagopal, A., Eaton, K., Meng, X., Plante, E., & Holland, S. K. (2012). A 10-year longitudinal fMRI study of narrative comprehension in children and adolescents. *NeuroImage, 63*(3), 1188-1195.

Szaflarski, J. P., Holland, S. K., Schmithorst, V. J., & Byars, A. W. (2006). fMRI study of language lateralization in children and adults. *Human Brain Mapping, 27*(3), 202-212.

Szaflarski, J. P., Rajagopal, A., Altaye, M., Byars, A. W., Jacola, L., Schmithorst, V. J., . . . Holland, S. K. (2012). Left-handedness and language lateralization in children. *Brain Res, 1433*, 85-97.

Szaflarski, J. P., Schmithorst, V. J., Altaye, M., Byars, A. W., Ret, J., Plante, E., & Holland, S. K. (2006). A longitudinal functional magnetic resonance imaging study of language development in children 5 to 11 years old. *Annals of Neurology, 59*(5), 796-807.

Tellez-Zenteno, J. F., Hernandez Ronquillo, L., Moien-Afshari, F., & Wiebe, S. (2010). Surgical outcomes in lesional and non-lesional epilepsy: a systematic review and meta-analysis. *Epilepsy Res, 89*(2-3), 310-318.

Terra-Bustamante, V. C., Inuzuca, L. M., Fernandes, R. M., Funayama, S., Escorsi-Rosset, S., Wichert-Ana, L., . . . Sakamoto, A. C. (2005). Temporal lobe epilepsy surgery in children and adolescents: clinical characteristics and post-surgical outcome. *Seizure, 14*(4), 274-281.

Thivard, L., Hombrouck, J., du Montcel, S. T., Delmaire, C., Cohen, L., Samson, S., . . . Lehericy, S. (2005). Productive and perceptive language reorganization in temporal lobe epilepsy. *NeuroImage, 24*(3), 841-851.

Thomason, M. E. (2009). Children in non-clinical functional magnetic resonance imaging (FMRI) studies give the scan experience a "thumbs up". *Am J Bioeth, 9*(1), 25-27.

Thornton, R. C., Rodionov, R., Laufs, H., Vulliemoz, S., Carmichael, D. W., McEvoy, A. W., . . . Lemieux, L. (2009). Exploring Haemodynamic Changes Linked to Focal Seizures Using Simultaneous Eeg-Fmri. *Epilepsia, 50*, 90-90.

Tie, Y., Suarez, R. O., Whalen, S., Radmanesh, A., Norton, I. H., & Golby, A. J. (2009). Comparison of blocked and event-related fMRI designs for pre-surgical language mapping. *NeuroImage, 47 Suppl 2*, T107-115.

Tourville, J. A., Reilly, K. J., & Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *NeuroImage, 39*(3), 1429-1443.

Tremblay, P., & Small, S. L. (2011). Motor response selection in overt sentence production: a functional MRI study. *Front Psychol, 2*, 253.

Turkeltaub, P. E., Gareau, L., Flowers, D. L., Zeffiro, T. A., & Eden, G. F. (2003). Development of neural mechanisms for reading. *Nat Neurosci, 6*(7), 767-773.

Turken, A. U., & Dronkers, N. F. (2011). The neural architecture of the language comprehension network: converging evidence from lesion and connectivity analyses. *Front Syst Neurosci, 5*, 1.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., . . . Joliot, M. (2002). Automated anatomical labeling of

activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage, 15*(1), 273-289.

Unser, M., Aldroubi, A., & Eden, M. (1993a). B-Spline Signal-Processing .1. Theory. *Ieee Transactions on Signal Processing, 41*(2), 821-833.

Unser, M., Aldroubi, A., & Eden, M. (1993b). B-Spline Signal-Processing .2. Efficient Design and Applications. *Ieee Transactions on Signal Processing, 41*(2), 834-848.

Vagharchakian, L., Vagharchakian, G., Dehaene Lambertz, C., Pallier, S., & Dehaene. (2012). A Temporal Bottleneck in the Language Comprehension Network. *The Journal of Neuroscience, 32*(26), 9089-9102.

Vannest, J., Rasmussen, J., Eaton, K. P., Patel, K., Schmithorst, V., Karunanayaka, P., . . . Holland, S. (2010). FMRI activation in language areas correlates with verb generation performance in children. *Neuropediatrics, 41*(5), 235-239.

Vannest, J. J., Karunanayaka, P. R., Altaye, M., Schmithorst, V. J., Plante, E. M., Eaton, K. J., . . . Holland, S. K. (2009). Comparison of fMRI data from passive listening and active-response story processing tasks in children. *J Magn Reson Imaging, 29*(4), 971-976.

Vigneau, M., Beaucousin, V., Herve, P. Y., Duffau, H., Crivello, F., Houde, O., . . . Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *NeuroImage, 30*(4), 1414-1432.

Visser, M., Jefferies, E., & Lambon Ralph, M. A. (2010). Semantic processing in the anterior temporal lobes: a meta-analysis of the functional neuroimaging literature. *J Cogn Neurosci, 22*(6), 1083-1094.

Vlooswijk, M. C., Jansen, J. F., Majoie, H. J., Hofman, P. A., de Krom, M. C., Aldenkamp, A. P., & Backes, W. H. (2010). Functional connectivity and language impairment in cryptogenic localization-related epilepsy. *Neurology, 75*(5), 395-402.

Vlooswijk, M. C., Vaessen, M. J., Jansen, J. F., de Krom, M. C., Majoie, H. J., Hofman, P. A., . . . Backes, W. H. (2011). Loss of network efficiency associated with cognitive decline in chronic epilepsy. *Neurology, 77*(10), 938-944.

Vulliemoz, S., Thornton, R., Rodionov, R., Carmichael, D. W., Guye, M., Lhatoo, S., . . . Lemieux, L. (2009). The spatio-temporal mapping of epileptic networks: Combination of EEG-fMRI and EEG source imaging. *NeuroImage, 46*(3), 834-843.

Wada, J. (1949). A new method for determination of the side of cerebral speech dominance. A preliminary report on the intra-carotid injectino of sodium amytal in man. *Igaku to Seibutasugaku, 27*, 1207-1219.

Waites, A. B., Briellmann, R. S., Saling, M. M., Abbott, D. F., & Jackson, G. D. (2006). Functional connectivity networks are disrupted in left temporal lobe epilepsy. *Annals of Neurology, 59*(2), 335-343.

Wang, J., Fan, L., Zhang, Y., Liu, Y., Jiang, D., Zhang, Y., . . . Jiang, T. (2012). Tractography-based parcellation of the human left inferior parietal lobule. *NeuroImage, 63*(2), 641-652.

Warren, J. E., Crinion, J. T., Lambon Ralph, M. A., & Wise, R. J. (2009). Anterior temporal lobe connectivity correlates with functional outcome after aphasic stroke. *Brain, 132*(Pt 12), 3428-3442.

Weber, B., Wellmer, J., Reuber, M., Mormann, F., Weis, S., Urbach, H., . . . Fernandez, G. (2006). Left hippocampal pathology is associated with atypical language lateralization in patients with focal epilepsy. *Brain, 129*(Pt 2), 346-351.

Wechsler, D. (1989). *Wechsler Preschool and Primary Scale of Intelligence – Revised*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1999). *Manual for the Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: The Psychological Corporation.

Weiller, C., Musso, M., Rijntjes, M., & Saur, D. (2009). Please don't underestimate the ventral pathway in language. *Trends Cogn Sci, 13*(9), 369-370; 370-361.

Wernicke, C. (1874). Der aphasische Symptomencomplex.

WHO. (2001). World Health Organisation: Epilepsy factsheet. from http://www.who.int/mediacentre/factsheets/fs999/en/

Wiig, E. H., Secord, W., & Messing Semel, E. (2004). *CELF Preschool 2: Clinical Evaluation of Language Fundamentals Preschool.* . San Antonio: Pearson/PsychologicalCorp.

Wilke, M. (2012). An alternative approach towards assessing and accounting for individual motion in fMRI timeseries. *NeuroImage, 59*(3), 2062-2072.

Wilke, M., Lidzba, K., & Krageloh-Mann, I. (2009). Combined functional and causal connectivity analyses of language networks in children: a feasibility study. *Brain Lang, 108*(1), 22-29.

Wilke, M., Lidzba, K., Staudt, M., Buchenau, K., Grodd, W., & Krageloh-Mann, I. (2005). Comprehensive language mapping in children, using functional magnetic resonance imaging: what's missing counts. *Neuroreport, 16*(9), 915-919.

Wilke, M., Pieper, T., Lindner, K., Dushe, T., Holthausen, H., & Krageloh-Mann, I. (2010). Why one task is not enough: functional MRI for atypical language organization in two children. *Eur J Paediatr Neurol, 14*(6), 474-478.

Wilke, M., & Schmithorst, V. J. (2006). A combined bootstrap/histogram analysis approach for computing a lateralization index from neuroimaging data. *NeuroImage, 33*(2), 522-530.

Wise, R. J. (2003). Language systems in normal and aphasic human subjects: functional imaging studies and inferences from animal studies. *Br Med Bull, 65*, 95-119.

Wood, A. G., Harvey, A. S., Wellard, R. M., Abbott, D. F., Anderson, V., Kean, M., . . . Jackson, G. D. (2004). Language cortex activation in normal children. *Neurology, 63*(6), 1035-1044.

Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J., & Evans, A. C. (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping, 4*(1), 58-73.

Wyllie, E., Comair, Y. G., Kotagal, P., Bulacio, J., Bingaman, W., & Ruggieri, P. (1998). Seizure outcome after epilepsy surgery in children and adolescents. *Ann Neurol, 44*(5), 740-748.

Yeatman, J. D., Ben-Shachar, M., Glover, G. H., & Feldman, H. M. (2010). Individual differences in auditory sentence comprehension in children: An exploratory

event-related functional magnetic resonance imaging investigation. *Brain and Language, 114*(2), 72-79.

Yerys, B. E., Jankowski, K. F., Shook, D., Rosenberger, L. R., Barnes, K. A., Berl, M. M., . . . Gaillard, W. D. (2009). The fMRI success rate of children and adolescents: typical development, epilepsy, attention deficit/hyperactivity disorder, and autism spectrum disorders. *Human Brain Mapping, 30*(10), 3426-3435.

Zsoter, A., Staudt, M., & Wilke, M. (2012). Identification of successful clinical fMRI sessions in children: an objective approach. *Neuropediatrics, 43*(5), 249-257.